DOCUMENT RESUME

ED 088 512 IR 000 346

AUTHOR	Collins, Allan; And Others
TITLE	Analysis and Synthesis of Tutorial Dialogues. Semi-Annual Technical Report Number Two.
INSTITUTION	Bolt, Beranek and Newman, Inc., Cambridge, Mass.
SPONS AGENCY	Advanced Research Projects Agency (DOD), Washington, D.C.; Office of Naval Research, Washington, D.C. Personnel and Training Research Programs Office.
REPORT NO	BBN-R-2789
PUB DATE	Mar 74
NOTE	92p.
EDRS PRICE	MF-$0.75 HC-$4.20
DESCRIPTORS	Artificial Intelligence; Computer Assisted Instruction; *Computer Programs; *Discourse Analysis; Educational Research; Geography Instruction; Higher Education; *Individualized Instruction; Information Processing; Interaction; Semantics; Tape Recordings; *Teaching Techniques; Thought Processes; Tutorial Programs; *Tutoring
IDENTIFIERS	Dialogue Analysis; *SCHOLAR; South America

ABSTRACT

An analysis was made of the strategies by which tutors adapt their teaching to individual students; the objective was to synthesize these strategies in a computer system called SCHOLAR. Tape recordings of dialogues between tutors and students about South American geography were made to discover the strategies used by tutors. Since SCHOLAR is a well-defined program, it was possible to analyze the naturalistic data in precise terms. Individual aspects of the dialogues were analyzed separately and, based upon the analyses, several hypotheses were prepared which described how the tutor related his teaching to the individual student. These strategies were implemented in SCHOLAR, and it seems likely that the analytical method employed could be extended to a wide range of conversational situations. For example, the use of this method (Dialogue Analysis) would permit psychologists to study questions about the interactive aspects of human language processing that cannot even be considered with traditional laboratory methods. (Author)

ED 088512

BOLT BERANEK AND NEWMAN INC

CONSULTING · DEVELOPMENT · RESEARCH

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: IR SP
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

BBN Report No. 2789

ANALYSIS AND SYNTHESIS OF TUTORIAL DIALOGUES

Allan Collins
Eleanor H. Warnock
Joseph J. Passafiume

Bolt Beranek and Newman Inc.
Cambridge, Massachusetts 02138

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

March 1974

Sponsored by

Office of Naval Research
and the Advanced Research Projects Agency

Approved for public release; distribution unlimited.
Reproduction in whole or in part is permitted for
any purpose of the United States Government.

IR 000 346

ED 088512

ANALYSIS AND SYNTHESIS OF TUTORIAL DIALOGUES[1]

Allan Collins

Eleanor H. Warnock

Joseph J. Passafiume

Bolt Beranek and Newman Inc.

Cambridge, Massachusetts 02138

Contract No. N00014-71-C-0228, dated March 1, 1971
Amendment Modification No. P00002, dated January 1, 1973
Expiration Date, June 30, 1974
Total Amount of Contract $436,096.00
Principal Investigator, Allan Collins [(617)491-1850]

Sponsored by: Office of Naval Research
Contract Authority No. NR 154-330
Scientific Officers: Dr. Marshall Farr and
Dr. Joseph Young

and

Advanced Research Projects Agency
ARPA Order No. 2284, dated August 30, 1973
Program Code No. 61101E

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency, the Office of Naval Research, or the U.S. Government.

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

Bolt Beranek and Newman Inc.

50 Moulton Street
Cambridge, MA 02138

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

ANALYSIS AND SYNTHESIS OF TUTORIAL DIALOGUES

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Semi-Annual Technical Report No. 2

5. AUTHOR(S) *(First name, middle initial, last name)*

Allan M. Collins, Eleanor H. Warnock, Josepn J. Passafiume

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
March 1974	78	18

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
N00014-71-C-0228	2789
b. PROJECT NO. NR 154-330	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Personnel and Training Research Programs, Office of Naval Research Arlington, VA 22217

13. ABSTRACT

In this paper we attempt to analyze the strategies by which tutors adapt their teaching to individual students, so that we can synthesize these strategies in a computer system called SCHOLAR. To find out what strategies tutors use, we tape-recorded dialogues between various tutors and students on the topic of South American geography. Because SCHOLAR is a well-defined program, it is possible to analyze such ill-defined naturalistic data in precise terms, with respect to the structure and processing of information in SCHOLAR. We analyzed the dialogues concentrating on one aspect at a time. Based on our analyses, we propose in this paper several hypotheses about how the tutor relates his teaching to the individual student. We show how in modified form we have implemented some of these strategies in SCHOLAR. We further argue that the analytical method employed here could be extended to a wide range of conversational situations. This method (Dialogue Analysis) would permit psychologists to study questions about the interactive aspects of human processing that cannot even be considered with traditional laboratory methods.

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Computer-assisted instruction						
Artificial intelligence						
Semantic networks						
Mixed-initiative computer systems						
Tutoring						

Abstract

In this paper we attempt to analyze the strategies by which tutors adapt their teaching to individual students, so that we can synthesize these strategies in a computer system called SCHOLAR. To find out what strategies tutors use, we tape-recorded dialogues between various tutors and students on the topic of South American geography. Because SCHOLAR is a well-defined program, it is possible to analyze such ill-defined naturalistic data in precise terms, with respect to the structure and processing of information in SCHOLAR. We analyzed the dialogues concentrating on one aspect at a time. Based on our analyses, we propose in this paper several hypotheses about how the tutor relates his teaching to the individual student. We show how in modified form we have implemented some of these strategies in SCHOLAR. We further argue that the analytical method employed here could be extended to a wide range of conversational situations. This method (Dialogue Analysis) would permit psychologists to study questions about the interactive aspects of human language processing that cannot even be considered with traditional laboratory methods.

TABLE OF CONTENTS

		Page
1.	INTRODUCTION.	1
2.	A BRIEF DESCRIPTION OF SCHOLAR.	2
3.	ANALYSIS OF TUTORIAL DIALOGUES.	10
	3.1 Methodology	10
	3.2 Tutorial Strategies	14
	Topic Selection	14
	The Interweaving of Questioning and Presentation	33
	Questioning about Basic Concepts.	42
	Reviewing by the Tutor.	44
	Providing Hints	47
	The Tutors' Response to Errors.	50
4.	SYNTHESIS OF TUTORIAL DIALOGUES	52
	4.1 Comparison of Computer and Human Dialogues. . .	53
	Topic Selection	54
	Interweaving of Questioning and Presentation. .	63
	Reviewing	64
	Error Correction.	66
	4.2 Use of SCHOLAR to Evaluate Teaching Methods . .	69
5.	DISCUSSION	70
	5.1 Comments on Dialogue Analysis as a Method . . .	70
	5.2 The Analytic-synthetic Approach	73
References		75
Footnotes.		77

1. INTRODUCTION

How does the tutor adapt his teaching to the individual student? This is like asking how a debater wins an argument or how a thief eludes a policeman: so much depends on what the other guy does that any experimental manipulation could distort the process beyond all recognition. We decided to study the question by recording actual dialogues between tutors and students, and playing these back over and over, each time listening to a different thread in the conversation.

There were two aspects of tutoring we initially wanted to look at. One ability of a human tutor is that he builds on what the student already knows. Because a tutor carries on a dialogue with the student, he can question him about his previous knowledge. Then he can teach new material by relating it to that previous knowledge.

Another ability of the tutor is that he can respond directly to student errors. If the student makes a mistake, the tutor can question him to diagnose the confusion and can provide relevant information to straighten it out. Hence, in the tutorial dialogue, the information taught can be directed against the existing confusions, rather than against what the teacher anticipates might be a typical student's confusion.

These two abilities of the tutor, coupled with the capability of responding to the student's questions, make the tutor's role worth modelling in a computer. We are attempting to develop such

interactive capabilities in a computer program called SCHOLAR so that, like the tutor, it can function in a responsive way to the student's knowledge, his misconceptions, and his questions.

Our approach has been to look first at the strategies that human tutors use in interacting with students. The analysis of tutorial strategies is akin to the protocol analysis of Newell and Simon (1972) or the analysis of children's speech by Brown and his colleagues (Brown, 1970, 1973). Based on this dialogue analysis, we are trying to approximate these strategies within SCHOLAR so that it can respond to each student as an individual.

2. A BRIEF DESCRIPTION OF SCHOLAR

In analyzing the tutorial dialogues we have used SCHOLAR (Carbonell, 1970a, 1970b, 1971) as a kind of filter. That is, we have selected those aspects of the dialogues that we can see how to program in SCHOLAR.

To understand SCHOLAR enough to follow the analysis, it is easiest to compare a small part of the data base with a dialogue between SCHOLAR and a student, based on that piece of data base. SCHOLAR's main data base is on South American geography (Figure 1). The program is basically independent of the subject matter, and in fact there is also a second data base on procedures for using the ARPA Computer Network. The data base is a semantic network (Quillian, 1968; Carbonell, 1970a; Collins and Quillian, 1972), which means that information about any concept is stored in terms of

```
CAPITAL
       [((CN CAPITAL))
        NIL
        (SUPERC NIL CITY (PLACE NIL (OF NIL GOVERNMENT)))
        (SUPERP (I 1) COUNTRY)
        (APPLIED\TO (I 4) COUNTRY STATE)
        (EXAMPLES (I 2)
                  ($EOR BUENOS\AIRES LIMA MONTEVIDEO BRASILIA
                        GEORGETOWN CARACAS BOGOTA QUITO SANTIAGO
                        ASUNCION LA\PAZ WASHINGTON]

BRASILIA
       [((XN BRASILIA BRAZILIA))
        NIL
        (SUPERC NIL CITY CAPITAL)
        (SUPERP (I 6 B) BRAZIL BRAZILIAN\HIGHLANDS)
        (POPULATION (I 1)
                    (APPROX NIL 400000))
        (LOCATION NIL (IN NIL (BRAZIL NIL CENTRAL))
                  (IN (I 1) BRAZILIAN\HIGHLANDS)
                  (LATITUDE (I 4) -16)
                  (LONGITUDE (I 5) -48]

BRAZIL
       [((XN BRAZIL BRASIL))
        NIL
        (SUPERC NIL COUNTRY)
        (SUPERP (I 6 B) SOUTH\AMERICA)
        (LOCATION NIL (IN NIL (SOUTH\AMERICA NIL NORTHEASTERN))
                  (LATITUDE (I 4) (RANGE NIL 5 -33))
                  (LONGITUDE (I 5) (RANGE NIL -35 -73))
                  (BORDERING\COUNTRIES (I 1)
                        (NORTHERN (I 1)
                                  ($EX GUYANA VENEZUELA SURINAM
                                       FRENCH\GUIANA))
                        (WESTERN (I 1)
                                 ($EX PERU COLOMBIA))
                        (SOUTHERN (I 1)
                                  ($EX URUGUAY BOLIVIA PARAGUAY
                                       ARGENTINA)))
                  (COAST (I 1)
                         (EASTERN (I 1) ATLANTIC)))
        (CAPITAL (I 1) BRASILIA)
        (CITIES (I 2)
                ($L SAO\PAULO RIO\DE\JANEIRO BRASILIA PORTO\ALEGRE
                    SALVADOR RECIFE SANTOS MANAOS BELEM]
```

Figure 1: Three Partial Entries from SCHOLAR's Geography Data Base.

other concepts entered in the data base.

The information is extensively cross-referenced. This can be seen by looking at the three entries shown: Brazil has Brasilia as a value for two different attributes, Cities and Capital. Brasilia has for its Superordinate (labeled SUPERC for Superconcept) two values, City and Capital. Brazil occurs under Brasilia as a value for two attributes, Superpart (labeled SUPERP) and Location. Capital has City as a Superordinate and Brasilia as an Example.

The entry for Location under Brazil illustrates another aspect of the data base, called embedding. Under the attribute Location there are several subattributes, among which is Bordering-Countries. But under Bordering-Countries there are the subattributes Northern, Western, and Southern, each of which has several values. Embedding describes the nesting of attributes or values, which can go down as deep as is necessary to describe a property completely.

In the data base there are also tags, such as the Nil after Location and the (I 1) after Bordering-Countries. These tags are called importance tags or irrelevancy tags (I-tags), and they vary from Nil, interpreted as (I Ø), up to (I 6). The lower the tag, the more important the piece of information. But they add up as you go down to lower embedded levels. Thus under Location in Brazil, Northeastern South America has an I-tag level of Ø, Bordering-Countries an I-tag level of 1, and each Northern Bordering-Country an I-tag level of 2. These I-tags are used in a number of ways by the program, as we will describe below.

In carrying on a dialogue with a student, SCHOLAR is guided in its questioning by an internal agenda. The agenda determines what topics are to be discussed and how much time is to be spent on each topic. However, the student can always interrupt with a question about any topic he chooses.

The agenda operates as a pushdown list. If the overall topic is South America, SCHOLAR must generate a subtopic, e.g., Brazil. In the version of SCHOLAR shown in Figure 2, the subtopics are generated on a random basis, weighted for importance in terms of I-tags. When a subtopic is selected, a portion of the total time is allocated to it. If the subtopic is still too general, as is Brazil, then a lower subtopic is generated in similar fashion, e.g., Topography of Brazil. All subtopics are generated automatically by SCHOLAR itself, except to the degree that the teacher or student wants to specify just what material should be covered. But generally SCHOLAR follows a dynamic agenda rather than a prespecified agenda.

Figure 2 shows a dialogue between SCHOLAR and a student. The student's answers and questions are enclosed in asterisks.

The student shown interacting with SCHOLAR is not a real student but the second author. The dialogue has been edited to leave out unessential parts. An actual dialogue with a high-school student is shown later on, in the section on Synthesis of Tutorial Dialogues. SCHOLAR has behaved fairly well in sessions with high-school students, but it still occasionally blunders, and it is

THE CITIES IN BRAZIL ARE ----.

RIO DE JANEIRO, BUENOS AIRES, AND TITICACA

"RIO DE JANEIRO" IS FINE.
I DON'T THINK BUENOS AIRES AND TITICACA ARE CORRECT.

YOU COULD ALSO MENTION SAO PAULO, BRASILIA, PORTO ALEGRE, SALVADOR, RECIFE, SANTOS, MANAOS, AND BELEM.

BUENOS AIRES IS NOT IN BRAZIL BUT IN ARGENTINA.

TITICACA IS NOT A CITY BUT A LAKE IN SOUTH AMERICA.

WHERE IS BRASILIA

BRASILIA IS IN CENTRAL BRAZIL.
IT IS IN THE BRAZILIAN HIGHLANDS.

WHAT IS THE MOST IMPORTANT DIFFERENCE BETWEEN BRASILIA AND RIO

THE PRIMARY DIFFERENCE IS:

BRASILIA IS IN CENTRAL BRAZIL.
BUT
RIO DE JANEIRO IS IN EASTERN BRAZIL.

SELECT AN ALTERNATIVE FROM THE LIST:

ASIA
NORTH AMERICA
AFRICA
SOUTH AMERICA

TO COMPLETE THE SENTENCE:

BRAZIL IS LOCATED IN ---.

BRAZIL IS IN SOUTH AMERICA

THAT'S RIGHT.

Figure 2a: Illustrative Dialogue between SCHOLAR and Student. (Student inputs are between asterisks.)

TELL ME ABOUT THE AMAZON

THE AMAZON IS NAVIGABLE.
IT IS A RIVER.
IT IS IN BRAZIL AND PERU.
THE LENGTH IS 4,000 MILES.

PLEASE TELL ME MORE ABOUT THE AMAZON

THE MOUTH IS IN THE ATLANTIC.
THE TRIBUTARIES ARE THE MARANON, THE RIO NEGRO, THE JURUA, THE PURUS, THE TAPAJOZ, THE MADEIRA, AND THE TOCANTINS.

IS SAO PAULO A CITY IN BRAZIL?

I THINK I HAVE ALREADY TOLD YOU ABOUT IT.

YES.

Figure 2b: Illustrative Dialogue between SCHOLAR and Student.

limited in that it can answer only certain types of questions that do not contain complicated syntactic constructions.

In the first exchange shown in the dialogue, SCHOLAR asked about the cities in Brazil. The student answered with three names, two of which were incorrect. Just because these two were not stored as cities in Brazil within the data base, though, did not mean that they were necessarily wrong. For example, Corumba is not stored, but it is a small city in Brazil. So SCHOLAR corrected the student by indicating which of her answers were not stored, but it hedged what it said. It also pointed out the difference between the student's answer and the correct answer, with respect to the Superordinate or Superpart. That is, it looked up Buenos Aires and found that it was in Argentina, not Brazil (wrong Superpart); it looked up Titicaca and found that it was a lake, not a city (wrong Superordinate); and it pointed out these differences.

Next the student asked where Brasilia was. Under the Location of Brasilia, the information with the lowest I-tags indicated that Brasilia was in central Brazil and in the Brazilian Highlands. Less important information about its longitude and latitude was omitted because the I-tags were too high, i.e., above the cutoff level used in answering most questions. But it could be retrieved by asking, for example, "What is the latitude of Brasilia?".

The student then asked about the most important difference between Brasilia and Rio de Janeiro. The subroutine that handled this question looked for common attributes under both Brasilia and

Rio de Janeiro in order of importance in terms of I-tags. It first checked the Superordinate, but the values were the same for both (i.e., City). Then it came to Location where the values were different and printed this out as the most important difference.

When the student stopped asking questions, SCHOLAR asked another question, chosen on a random basis weighted for importance in terms of I-tags. It was a multiple-choice question about the location of Brazil, which the student answered correctly. The three wrong choices were generated by picking other examples from the Superordinate of the correct answer, South America (in this case other continents).

The next two questions by the student illustrate another use of the I-tags. The first of the two, "Tell me about the Amazon", produced information down to a level of 1 with respect to I-tags. The second question, "Tell me more about the Amazon", produced all the information from level 2 through level 3 with respect to I-tags. It would be possible to go on asking for more, and getting information that was less and less relevant.

The final question by the student, about Sao Paulo, referred to information SCHOLAR had told her earlier. At that time, SCHOLAR had left a temporary tag in the data base on this piece of information. Then, when the student raised the question again, SCHOLAR recognized that they had discussed it previously, and pointed that out.

This dialogue is intended to give the flavor of the SCHOLAR

system. The program itself is described in much more detail in Carbonell (1970a, 1970b, 1971). In Section 4 we will describe the tutorial version of SCHOLAR, which incorporates what was learned from the analysis of human tutorial dialogues. Here, we have tried to give enough background to understand the dialogue analysis described in the next section.

3. ANALYSIS OF TUTORIAL DIALOGUES

3.1 Methodology

In order to determine in general terms what is involved in modelling a tutor's teaching strategy, we collected tape recordings of four tutors discussing South American geography with several different students. In all, we tape-recorded ten dialogues with different combinations of the four tutors and six students. The principal tutors (RP and AC, the first author) each tutored in four of the dialogues. Both have extensive teaching experience at the college level, though neither has taught geography. The third and fourth tutors each taught only one session, and did not prepare nearly as extensively as the first two tutors. The students were employees at BBN. Two of the students had read material on South American geography and the rest had not. The students varied widely in their sophistication about geography. In all the sessions, the tutors could point out different places on an unlabeled map of South America. In most of the sessions, the tutors conducted the dialogues in the manner that they thought would be most effective.

However, in two cases, the tutors tried to limit themselves (like SCHOLAR but not too successfully) to questions that would evoke specific answers, such as names, or lists of things. Because of the differences in tutorial method, in student preparation, and in student sophistication, the dialogues varied widely.

We did not attempt to analyze a large sample that could be generalized to most tutors and students but instead tried to analyze a few tutors in depth. Our aim was to assess the individual tutorial strategies of people we considered to be good teachers. It was necessary to study a tutor with several different students in order to be able to abstract the commonalities in his approach. One of the tutors was also the first author of this paper. This was done because it is a little easier to analyze what processing a tutor must have gone through to arrive at a given output, given that something is known about what he has stored and what he does not have stored. However, the analysis in this paper centers on the strategies used by the first tutor, and the strategies of the other tutors are presented primarily for comparison.

Natural data of this kind are assiduously avoided by psychologists because of the difficulty in analyzing what is collected. But as Newell and Simon (1972) have shown with protocol analysis, it is possible to analyze such data in terms of a computer model. With SCHOLAR as a model, we can analyze different aspects of the dialogues by specifying how they could be produced by SCHOLAR or by certain procedures added to SCHOLAR.

The technique of protocol analysis has been developed and extensively used by Newell and Simon to create information processing models of human problem solving and of thought processes in general. Our use of tutorial dialogue analysis is distinct from their work in three ways. First, we do not perform an exhaustive analysis of the dialogue. Rather, we study one particular aspect at a time, throughout the dialogue (e.g., error correction strategies, the use of hints, the selection of topics and subtopics, etc.). Second, in protocol analysis, a person is usually performing a non-verbal task (e.g., playing chess) and being forced to verbalize his thinking. By contrast, in our dialogue analysis, the verbalization is inherent in the task of carrying on a dialogue, and recording does not create any interference with the task. Third, there is more of an inferential process in our analysis of dialogues than in their analysis of protocols. Because the tutorial strategies are not verbalized by the tutor, they must be inferred from what he says. Our inferences are certainly as prone to error as the reader's, and there is no definitive evidence that our hypotheses are the correct ones.

There are two points related to our methodology worth making here, because of the inferential nature of such analysis. First, we will present samples of the raw data from which we have derived our descriptions so that the reader may compare our descriptions with his own analysis. Second, since our descriptions of tutorial strategies are realizable in SCHOLAR or in other computer programs, our hypotheses can be tested in terms of how well they can produce

new dialogues that look like tutor-student dialogues. This, of course, is Turing's (1950) test for evaluating computer programs. It is also in essence the same procedure used to evaluate any psychological model. That is, the test comes down to a comparison between the model's prediction (the computer output) and the human data actually obtained.

We should point out that our psychological description may be more specific or less specific, yet still be well defined. For example, a fairly specific description of a tutor's behavior might be that he asks a "where" question in situation X, and a "what" question in situation Y. But often there is not enough data to specify situations X and Y, so it is a more correct (and less specific) description to say that he asks a "what" or a "where" question in situation Z, which includes situations X and Y. This is similar to the rounding-off problem in arithmetic where one does not want to present more significant digits in the answer than the calculation procedure permits. There is another reason for not being too specific: details about when the tutor asks a "where" question as opposed to a "what" question are probably not very important with respect to how well the tutor teaches or the student learns. In selecting only a few top-level aspects of the dialogues, we concentrated on what we think are the essentials of a good tutorial strategy. But because details must be specified for any computer implementation, we may have erred in the direction of more specificity than is justified.

It should be emphasized that in undertaking this experiment we had little idea of what we might find. Basically, we only had a framework in terms of SCHOLAR for analyzing the dialogues. In this sense, SCHOLAR acted as an information-processing model of the tutor. The need for synthesis in the program directed our analysis, and the results of the analysis, of course, determined our synthesis in the SCHOLAR program.

3.2 Tutorial Strategies

There were a number of different aspects of the way the tutors ran the dialogues in order to tailor the discussion to the individual student. We isolated six aspects, each of which we discuss below: topic selection, the interweaving of questioning and presentation, questioning about basic concepts, reviewing by the tutor, the use of hints, and the tutor's response to errors.

As suggested in the Introduction, we were initially interested in how the tutor related his teaching to the student's prior knowledge (which is discussed in the section on the Interweaving of Questioning and Presentation) and in how the tutor corrected student errors. The other aspects occurred to us as we studied the dialogues.

Topic Selection. The tutors appeared to generate topics as they went along rather than following a prespecified agenda. The topics taken up by the tutors in the dialogues were clearly not generated on a random basis. In some cases the topics were selected

by moving from place to place on the map. But more generally topic selection seemed to follow the structure of the tutor's knowledge about South American geography. This showed up in the fact that the pattern of topics selected looked rather like an outline for a course.

Figure 3 shows half of a dialogue with the first tutor. The organization of the dialogue can be seen in the outline of topics derived from the dialogue, shown in Figure 4. We will compare this organization to the organization found in other dialogues involving both this tutor and other tutors. Then we will describe our hypothesis as to how tutors select topics.

A tutor generally indicates the organization he is following by his comments when he introduces a new topic. In the dialogue shown in Figure 3, the first sentence established the overall topic, South America, and the third sentence introduced the first subtopic, Names-and-Locations-of-Countries[2]. This subtopic came up first in almost all the dialogues, probably because it is the most important property of South America.

The second subtopic in the dialogue shown was introduced by the two sentences in line 35: "That covers the major countries of South America. How about some of the geographical features?" Under the attribute Geographical-Features the tutor and student discussed Cape Horn, the Andes, and the three major river systems. The topics that

1. T. We are going to talk about South America today. Now, in order to begin, it would be helpful if I had some idea of what you know about South America to start with. So I wonder if you can tell, for example, do you know the name of that country? (points to Brazil)

2. S. Brazil.

3. T. How about this one? (points to Argentina)

4. S. Argentina.

5. T. Can you name the rest of them?

6. S. Chile. (points to Chile) Peru. (points to Peru) The Guianas. (points to the Guianas)

7. T. Do you know which ones are which?

8. S. No.

9. T. OK, we'll work on that. How about this one? (points to Venezuela) (not heard by student)

10. S. Is this Ecuador? (points to Colombia)

11. T. Which one? This one? (points to Ecuador)

12. S. Right here. (points to Colombia)

13. T. No, that's not Ecuador.

14. S. Oh dear!

15. T. Ecuador is near there. This is Ecuador. (points to Ecuador) Ecuador is much smaller than this country. (error correction) That is the one that connects to Panama. (hint)

16. S. No.

17. T. That's Colombia.

18. S. Colombia.

19. T. OK, and this is Ecuador. (points again to Ecuador) That leaves a few.

20. S. Is this Uruguay? (points to Paraguay)

Figure 3a: A Dialogue between the First Tutor and a Student[3]

21. T. No.

22. S. Is it Paraguay?

23. T. This is Paraguay. (points to Paraguay)

24. S. That has to be Uruguay. (points)

25. T. And that's Uruguay. (points)

26. S. And Venezuela? (points to Bolivia) (not heard)

27. T. You can remember Uruguay because it's so close to Argentina and sort of part way between Brazil and Argentina on the coast. (error correction)

28. S. Oh.

29. T. OK, that means you don't know what this one is? (points to Bolivia)

30. S. Is that Venezuela?

31. T. No, that's Bolivia. (error correction)

32. S. Bolivia.

33. T. Now, we've got everything except Venezuela and there must only be one left. (points to Venezuela)

34. S. I never would have guessed. (ironic)

35. T. That covers the major countries of South America. How about some of the geographical features. What things do you remember most about the layout of South America?

36. S. The Cape. Is that Cape Horn? (points to Cape Horn)

37. T. That's Cape Horn. Right at the base of South America is Cape Horn. Right about there. (points to tip)

38. S. Where all the shipping used to go around there, many years ago.

39. T. That's right, and some of the explorers. Remember the names of any of the explorers?

40. S. H'm. Magellan.

Figure 3b: A Dialogue between the First Tutor and a Student

41. T. There's even something in between there, called the Straits of Magellan.

42. S. Balboa. (ignored by tutor because didn't hear)

43. T. Down in there somewhere. (points to Straits)

44. S. Right down there.

45. T. OK. How about rivers or mountains, or mountain ranges?

46. S. Amazon.

47. T. The Amazon. Where do you think that is?

48. S. It's in here. (points to mouth)

49. T. Right, that's the Amazon. This map isn't too good, but the Amazon penetrates all the way back into here (points into Peru, Colombia) and there's lots of feeders.

50. S. Tributaries.

51. T. Tributaries or feeders into the Amazon. And it's the biggest river in terms of volume in the world. OK. How about some other land features that might be of interest?

52. S. The mountains in through here. (moves along upper coast of Pacific)

53. T. In through there? (pointing)

54. S. H'm An...(tries to think of name)

55. T. You're close.

56. S. Andes.

57. T. Yeah, the Andes. The Andes run all the way down and they sort of spread out here (points to Bolivia) and into Chile. Down into Chile.

58. S. Oh, into Chile.

59. T. So the Andes sort of define that coast almost. And they have some of the highest mountains in the world. And one of the highest mountains in the world is in Argentina.

Figure 3c: A Dialogue between the First Tutor and a Student

60. S. In Argentina? (echoic)

61. T. It's called Aconcagua.

62. S. I've never heard of it.

63. T. It's the highest mountain in North and South America. It's in the Argentine Andes. It's right there. (points)

64. S. That means "with water." (In Spanish "con" means "with" and "agua" means "water", but her inference is wrong. The tutor probably did not understand, and so ignored it.)

65. T. Now, an interesting dividing point between Uruguay and Argentina is the Uruquay River which is this river running right down here. (points to it) Then there's also another river.

66. S. That's right on the border of Argentina.

67. T. On the border of Argentina and Uruguay.

68. S. And Brazil.

69. T. Yes. It forms the border with Brazil as well as the border with Uruquay down here. And there's another river off here, which sort of all come together in the estuary. This is the Parana River.

70. S. The Parana River.

71. T. And those two come together in an estuary. Do you know what an estuary is?

72. S. Well, I think it's the mouth, isn't it?

73. T. Like a mouth. (echoic)

74. S. A mouth of a river or a bay. It flows out into the ocean.

75. T. How many different kinds of mouths of rivers are there?

76. S. Hmm

77. T. For example, what about the mouth of this river. (pointing to Orinoco) Is that an estuary?

78. S. Is it an inlet? (probably still talking about an estuary)

79. T. The kind of mouth with sort of a series of pieces coming out.

Figure 3d: A Dialogue between the First Tutor and a Student

80. S. Is an estuary like little tiny islands, perhaps in the middle of it?

81. T. No, an estuary is a kind of an open area where the current comes out into the ocean.

82. S. Into the ocean.

83. T. So that you have the mixing of the salt and fresh water. Where the tide and the current come together.

84. S. So it backs up into the river itself.

85. T. Whereas this kind of a mouth (pointing to Orinoco) is called a delta, where it sort of filters through. It's a branching of the river at the foot, built up by the silt that's carried down the river. The other large river in South America is the Orinoco. It's right here.

86. S. In Venezuela.

87. T. Yeah. The Orinoco has a well-defined delta, better defined than the Amazon.

88. S. Now a delta is what? Where the sand shifts?

89. T. Where the silt from the river has washed down to the mouth of the river and built up a little island-like area. And then the water sort of washes through the set of islands.

90. S. It almost forces an island into being.

91. T. In a way, between the fingers of the delta, there are islands.

92. S. Little flat lands.

93. T. Yeah, they're flat islands and very fertile. Now, one thing you probably haven't been exposed to is the climate and the regions of South America from the point of view of the nature of the terrain and those kinds of things. First, there's the equator. Do you know where the equator runs?

94. S. Right here. (indicating a line about Uruguay)

95. T. No. It's actually quite a bit north of that. (error correction)

96. S. Is it?

Figure 3e: A Dialoque between the First Tutor and a Student

97. T. The equator is way up here. (pointing along it) The equator runs right through the mouth of the Amazon.

98. S. That's higher than I thought.

99. T. And the latitude roughly equivalent to New York is somewhere down here. (pointing around Buenos Aires) Our kind of latitudes are pretty far down. So the area particularly around the Amazon is very tropical jungle. And that sort of encompasses the whole region in here where you have typical tropics. (indicates regions of tropics)

100. S. Rain forests.

101. T. And rain forests.

102. S. Lizards. Flowers and shrubbery.

103. T. And very rich ... Very rich dense tropical growth. Then an area that seems to be quite usable or farmable is the region down toward here (pointing to Pampas), which is called the Pampas. (spells) That's tillable farm land.

104. S. That's right where the estuary is.

105. T. Yes, it's down in that region. Then most of Brazil turns out to be treeless rolling grasslands, called a savanna.

106. S. A savanna.

107. T. Have you ever heard of that term before?

108. S. No.

109. T. A savanna is an area that we don't have in our country. It has lots of rain in the summertime and virtually no rain in the wintertime. While it's possible for people to live there, they really need to have some control of their water supply. If you want to farm it, you have to do something about water in the wintertime.

110. S. It rains more in the winter than in the summer.

111. T. It rains from more or less October to April, which is the warm season.

112. S. Which is their summer!

Figure 3f: A Dialogue between the First Tutor and a Student

113. T. Right. October to April.

114. S. And in the wintertime, it doesn't snow or anything. Too hot.

115. T. And it's too dry. Maybe now we'll take up the countries one at a time. Let's think about Argentina for a while. While we're working on Argentina, let's review and see if we can remember the countries that surround it.

116. S. All right.

117. T. What would this country be? (points to Chile) (review)

118. S. Chile.

119. T. Chile. (then he points to Bolivia) (review)

120. S. That's Bolivia.

121. T. Good. (points to Paraguay) (review)

122. S. Paraguay.

123. T. Paraguay. (points to Uruguay) (review)

124. S. And that's Uruguay.

125. T. Uruguay, and this one? (points to Brazil) (review)

126. S. That's Brazil and this is the Uruguay River.

127. T. Good.

128. S. And then this is the Parana river.

129. T. Gee. Very good.

130. s. And there's an estuary there.

131. T. And there's an estuary there. And there's an island off the tip of what cape? (review)

132. S. Cape Horn.

133. T. And the name of that island? Do you know that?

134. S. No.

135. T. That's Tierra del Fuego.

Figure 3g: A Dialogue between the First Tutor and a Student

136. S. Oh, Tierra del Fuego.

137. T. And the northern part of Argentina has a large sort of semi-arid plain that extends into Paraguay. And that's a plains area that is relatively unpopulated.

138. S. Why?

139. T. Because it's pretty dry.

140. S. In other words, it has the same problem as the savanna.

141. T. Yeah. That's called the Chaco (spells) and it extends all the way up into Paraguay and Bolivia.

142. S. Chaco.

143. T. Like most South American countries, the religion of Argentina, in fact probably the religion of 90% of South America is Catholic. And it is Spanish-speaking, and Argentina is also Spanish-speaking. There's one country that is not Spanish-speaking. Do you know what that one is?

144. S. British Guiana?

145. T. Well, the Guianas are a good exception, but excluding the Guianas, there's still a country.

146. S. OK, there's still a country that doesn't speak Spanish.

147. T. Yeah. (waits 7 secs and gives up) Brazil speaks Portuguese and actually was originally a Portuguese colony.

148. S. Oh, I didn't know that.

149. T. Well, the climate of Argentina is temperate, because of the fact that its latitude is down here.

150. S. The same as ours.

151. T. Do you know what a temperate climate is?

152. S. Where the average temperature would fall around 60. Is it something like that?

153. T. Mmhm, and it has a warm and a cold season.

(The dialogue went on to cover the other countries one at a time.)

Figure 3h: A Dialogue between the First Tutor and a Student

the first tutor brought up under Geographical-Features of South America in two other dialogues partially overlapped those in the dialogue shown. In one case he took up the two major lakes, in addition to the Andes and the three major river systems; in the other case he talked only about the three major river systems. In the latter case he probably forgot to bring up the Andes. Shortly thereafter, when he had taken up Argentina as a topic and was discussing the various regions of Argentina, he brought up the Andes as one of the regions. The Andes then apparently replaced Argentina as a new topic, and he discussed the entire Andes chain, rather than just the part of the Andes in Argentina. The variation in the subtopics discussed under Geographical-Features is evidence for the notion that the subtopics covered are not prespecified as part of a plan, but rather are facts stored under the topic in the tutor's data base. The only plan seems to be to cover the most important information stored under each topic.

The first tutor took up both topics mentioned above (i.e., Names-and-Locations-of-Countries and Geographical-Features) in three of his four dialogues. The other top-level attributes which the first tutor brought up under South America in different dialogues were: Climate-and-Terrain-in-Different-Regions (once), Minerals-of-South-America (twice), and Population-Characteristics (once). We think that the topics selected reveal the structure of his information about South America, and that the frequency with which they were selected reflects their relative importance in his mind.

It is interesting to note that all these topics also came up as subtopics in discussing specific countries (except that Names-and-Locations-of-Countries was changed to Bordering-Countries). See, for example, the outline in Figure 4 where Geographical-Features occurs as a topic under Argentina. This would indicate that the structure of this tutor's data base at the level of individual countries was parallel in part to the structure at the level of South America.

The second tutor showed a slightly different structure of information about South America. In contrast to the variety of topics that the first tutor discussed as top-level attributes under South America, the second tutor apparently had only two major divisions of information under South America: The first division might be called Geophysical-Information, and included names of bodies of water, names of countries, the location of the equator and the extent of the tropic zone, and what he called Land-Features, namely the mountains and rivers. The second division he referred to as Geopolitical-Information, and included information about the population, governments, and history of South America.

The other two tutors, each of whom ran one session, showed less organization of South American geography than the first two tutors. This is probably because they spent less time beforehand organizing the information in their own minds. Part of what a tutor must do to prepare (if we may speculate) is to create a number of intermediate concepts like Geographical-Features in his data base, under which he

groups the various facts he has stored. Hence, the less his preparation, the less structure there will be among the topics he discusses with the student.

The tutor discusses information under the current topic or subtopic mostly in order of importance. When he has exhausted all the important information under a subtopic, he pops back up to the previous topic. (It is this strategy that produces the outline.) For example, in the dialogue shown, the tutor popped back up to Geographical-Features several different times, in lines 45, 51, 65 and 85 of Figure 3. The pop-up is usually accompanied by a pause for thinking, where the tutor might say "O.K." or "now", as he did in three of the pop-ups (lines 45, 51, and 65) to Geographical-Features.

Though the tutor picks topics mainly according to importance, context influences his selection in two ways. An answer given by the student or a piece of information presented by the tutor may become the new topic for a period of time. For example, when the student answered Cape Horn in line 36 to the question about Geographical-Features, Cape Horn became the topic for several minutes of discussion. If the student had not named Cape Horn, the tutor probably would not have mentioned it at all under this topic. He did not bring it up in the other two dialogues in which he discussed Geographical-Features of South America. We think he expected the Amazon or the Andes as an answer, because after discussing Cape Horn he rephrased the question about

Geographical-Features (line 45) in terms of rivers or mountains. Thus the discussion of Cape Horn took place not because the tutor regarded it as the most important Geographical-Feature of South America, but because the student brought it up in the context of Geographical-Features.

But even when the student gives an expected answer, as with the Amazon and the Andes in lines 46 and 56, or when the tutor himself introduces a topic, as with Aconcagua, the Uruguay River, or the Orinoco River in lines 61, 65, and 85, it usually becomes the subtopic for a while. When this happens, properties of these subtopics come up, such as the fact that the Amazon has a large volume, that are less important than things discussed subsequently, such as the fact that the major mountain range is called the Andes. Hence going deeper into each of these subtopics for a period of time distorts the process from taking up topics purely in the order of their importance. This is the major way that context affects the discussion.

Every property that comes up in discussing one topic can in this way potentially provide the next topic, and topics may thus follow contextually related sequences. This method of selecting topics exhibits all the aspects of tracing paths in a semantic network (Quillian, 1968).

But tutors do not follow contextual sequences endlessly. When context leads the discussion to topics that are not very important, tutors pop up to more important topics. Context does, however, tend

to dominate selection of topics in this way where the material is not highly structured, as with the third and fourth tutors. Because they did not prepare and thus did not create a structure of intermediate concepts for grouping related facts (i.e., intermediate concepts such as Geographical-Features or Geopolitical-Information), they did not show this pattern of popping-up out of context to more important topics as much as the first two tutors did.

Context also affects the tutor's choice when he is popping up to select a new topic. In this situation he tends to pick a new topic that is related to a previous topic. For instance, when the tutor was talking about a major river system (lines 69-84) that included the Rio de la Plata (an estuary), he was led into a discussion of the difference between estuaries and deltas. During this discussion he pointed out (without naming the river) the particularly large delta of the Orinoco River. Then when he popped up out of the Rio de la Plata system to select a new topic in line 85, he selected the Orinoco. This selection was almost inevitable after its delta had been discussed. Another clear example occurred in a different dialogue where he was discussing Bolivia and mentioned tin as its major source of income. When he popped up out of Bolivia as a topic, he selected Minerals-of-South-America as the new topic for several minutes. Two more examples of context biasing the selection of a new topic after popping up are pointed out later in this section.

There is a variation on this kind of contextual influence.

When the tutor is discussing individual countries one at a time (as did the first tutor at the end of the dialogue shown), he goes from country to country, each time picking a neighboring country that has not yet been discussed. This is a kind of map-guided contextual selection, where relative importance is overridden by context. Tutors seem never to go from country to country in anything like their order of importance. A map-guided strategy can also be used to select Geographical-Features of South America. The fourth tutor, for example, started following along the coast from Cape Horn to the Rio de la Plata, and then to the mouth of the Amazon, and so on.

The selection strategy we have described here is relatively easy to formalize, though it is impossible to predict which topics will be selected unless the current state of the data base is known perfectly. The selection is also at the mercy of whatever topics the student raises. In its simplest form it can be described by the following set of rules, which are to be applied cyclically:

1. When the topic is an attribute (e.g., Geographical-Features), select the most important unused value under the current topic. When the topic is a value (e.g., South America or Cape Horn), select the most important attribute and value under the current topic. (Context affects this selection by temporarily increasing the importance of topics that are related to the previous topics discussed.)

2. If the attribute and value selected are below some criterion level of importance, which indicates that all the important information under the current topic has been exhausted, then pop up from the current topic to the previous topic in the pushdown list of topics, and start again at Rule 1. (The criterion level appears to depend on some combination of importance weighted by the time available. Factors affecting the criterion are discussed in the next section.)

3. The attribute and value selected are above the required level of importance, so formulate a question about the value of the attribute, or present the attribute and value to the student. (What determines whether there is a question or a presentation is discussed in the next section.)

4. Add new topics to the pushdown list of topics. (This is the major way context affects the selection of topics.) When the current topic is an attribute, the new value is added to the top of the pushdown list. When the topic is a value, first the new attribute and then the new value are added[4]. If the student gives an unexpected correct answer, his value is used instead of the value from the data base in adding to the pushdown list. If an answer is incorrect, an error correction strategy, discussed later, takes over temporarily.

5. The top item on the pushdown list of topics becomes the next topic.

Let us briefly explain how these five rules would operate to account for the topics selected by the first tutor in a portion (lines 35-45) of the dialogue in Figure 3. In line 35 the tutor popped up by Rule 2 from the subtopic Names-and-Locations-of-Countries to the top-level topic South America. By application of Rule 1, where the topic is a value (South America), the tutor selected Geographical-Features as the attribute, and probably either the Amazon or Andes as the value. The attribute and value selected were quite important, so no pop-up occurred by Rule 2. Then using Rule 3, a question was formulated about the property selected ("How about some of the geographical features?"). When the student gave an unexpected correct answer, Cape Horn, it replaced the Amazon or Andes, as prescribed by Rule 4. At that time both Geographical-Features and Cape Horn were entered on the pushdown list of topics by application of Rule 4 and Cape Horn became the new topic by Rule 5.

We will pursue the subsequent cycles through the five rules in slightly less detail. Applying Rule 1 with Cape Horn as topic, the tutor selected the attribute Location and the value At-the-Base-of-South-America. Location was probably selected because the student had mentioned the location of Cape Horn in her answer or because location was the most important property of Cape

Horn. After the information about the location was presented by Rule 3, both attribute and value were added to the pushdown list of topics by Rule 4. Other information about the Base-of-South-America or about the Location-of-Cape-Horn was either not worth discussing or not stored, so by Rule 2 the topic popped up again to Cape Horn. The next property selected by Rule 1 was the attribute Explorers-Who-Sailed-Past-Cape-Horn and the value Magellan; the choice probably was affected by context, in particular by the student's mention of shipping around Cape Horn. Applying Rule 3 the tutor formulated a question about the property ("Remember the names of any explorers?"), to which the student gave the expected answer, Magellan. At this point, the tutor might have discussed Magellan or other explorers who sailed past Cape Horn, such as Drake, but he did not. He popped up from both these topics by Rule 2, back to Cape Horn. Applying Rule 1 again with Cape Horn as the topic, he next selected a Geographical-Feature of Cape Horn called the Straits of Magellan. Here again the selection in Rule 1 appears to have been influenced by the previous mention of Magellan, since the tutor picked the Straits of Magellan rather than the more important Geographical-Feature called Tierra del Fuego, which was the one he selected later when he returned to Cape Horn in lines 131-136. After presenting (by Rule 3) the Location-of-the-Straits-of-Magellan, there was a series of pop-ups by Rule 2 through the Straits of Magellan, Geographical-Features-of-Cape-Horn, and Cape Horn itself back up to Geographical-Features-of-South-America. The tutor probably did not know any more about the Straits of Magellan,

but he might conceivably have popped back to one of the other two topics. For example, he might have mentioned Tierra del Fuego, or other facts about Cape Horn, such as its proximity to Antarctica. He presumably rejected these topics by Rule 2 as not important enough, though there may be some overriding mechanism that governs when to pop up in this way, such as using up too much time on Cape Horn as a topic.

The five rules will produce an outline of topics very much like that shown in the Figure 4, given a highly structured data base. For a data base that lacks a hierarchical structure in which the tutor can continually pop up, these rules will produce a wandering discussion, because context will dominate the selection of topics.

We do not argue that these rules describe perfectly how the tutor selects topics. For example, the rules in unmodified form cannot handle a case like the one where the first tutor took up the entire Andes under Argentina, after he had forgotten to mention them earlier under Geograhical-Features. Also, the tutor may forget what a previous topic was, because it gets too far down in his pushdown list. But we would argue that these rules, or something like them, are the predominant determinants of topic selection.

The Interweaving of Questioning and Presentation. One of our a priori questions about the dialogues was how tutors would combine the questioning of the student and the presentation to him of new material. As should be apparent from the dialogue in Figure 3, the tutor does not simply ask questions first to find out what the

```
South America
    Names and locations of countries
        Brazil (Q), ..., Venezuela (P)
    Geographical features
        Cape Horn (Q)
            Location
                Southern tip (Q)
            Explorers who sailed past Cape Horn
                Magellan (Q)
            Geographical feature
                Straits of Magellan (P)
        Rivers
            Amazon (Q)
                Location and extent
                    Brazil (Q)
                    Peru and Colombia (P)
                    Tributaries to the south (P)
                Volume
                    Largest in the world (P)
        Mountains
            Andes (Q)
                Location and extent
                    Venezuela (P), ..., Chile (P)
                Shape
                    Same as coastline (P)
                Height
                    Some of highest in the world (P)
                Peaks
                    Aconcagua (P)
                        Height
                            Highest in Americas (P)
                        Location
                            Argentina, at mid-point (P)
        Rivers
            Rio de la Plata system (P)
                Subparts
                    Uruguay River (P)
                        Location
                            Between Uruguay, Argentina, and
                            Brazil (P)
                    Parana River (P)
                        Location
                            Argentina (P)
```

Figure 4a: An Inferred Outline of Topics for the Dialogue of Figure 3 in Quasi Attribute-Value Form[5].

```
                    Estuary (i.e., Rio de la Plata) (P)
                        Definition (or Superconcept)
                            Mouth of river that flows into ocean (Q)
                        Contrast to (or Opposite of)
                            Delta (QP)
                                Definition
                                    Mouth with little islands (QP)
                        Definition (repeated as error correction)
                            Open area with mixing of salt and fresh
                             water (P)
            Orinoco (P)
                Location
                    Venezuela (P)
                Mouth
                    Delta (P)
                        Definition (repeated at request of student)
                            Mouth where silt has made islands (P)
Climate and terrain in different regions
    Equator (P)
        Location
            At mouth of Amazon (QP)
    Latitude equivalent to New York (P)
        Location
            Around Buenos Aires (P)
    Amazon region (P)
        Climate and terrain
            Tropical jungle (P)
        Extent
            Parts of Brazil, etc. (P)
    Pampas
        Location
            Around Buenos Aires (P)
        Use
            Farming (P)
    Brazilian Savanna (Mato Grosso) (P)
        Terrain (or Definition)
            Treeless rolling grassland (QP)
        Climate
            Rain in summer, no rain in winter
        Uses
            Farming, if control water supply
        Climate (repeated at request of student)
            Rain in summer (P)
            No precipitation in winter (S)
            Hot in winter (S)
            Dry in winter (P)
```

Figure 4b: An Inferred Outline of Topics for the Dialogue of Figure 3 in Quasi Attribute-Value Form.

```
Countries
    Argentina (P)
        Bordering countries (review by tutor)
            Chile (Q), ..., Brazil (Q)
        Geographical features
            Rio de la Plata System (S)
                Subparts
                    Uruguay River, Parana, Estuary (S)
            Cape Horn (Q)
                Island
                    Tierra del Fuego (QP)
            Chaco (P)
                Location and extent
                    Northern Argentina, Paraguay (P)
                Terrain
                    Semi-arrid plain (P)
                Population density
                    Relatively unpopulated (P)
                Extent (reiterated by tutor)
                    Into Paraguay and Bolivia (P)
        Population
            Religion
                90% Catholic (P)
            Language
                Spanish (P)
                    Exception to Spanish-speaking
                        Guianas (Q)
                        Brazil (QP)
                            Language
                                Portuguese (P)
        Climate
            Temperate (P)
                Because of
                    Latitude (P)
                Definition
                    Average temperature around 60° (Q)
                    Warm season and cold season (QP)
```

Figure 4c: An Inferred Outline of Topics for the Dialogue of Figure 3 in Quasi Attribute-Value Form.

student knows and then present new material that the student has not included in his answers. Nor does the tutor follow the "programmed learning" strategy of presenting some information, then asking questions about it, and then presenting some more information and asking questions about that, etc. Instead, there is an intricate interweaving of question and presentation that is tied to the structure of the topics that are selected.

One striking fact about the tutor's questions is that they often occurred when the tutor had popped back up from a lower level. If the outline of topics in Figure 4 is compared to the corresponding dialogue, there is a consistent pattern as to where the questions occurred. To facilitate the comparison, the topics that were brought up as questions are indicated by a Q in the outline, and those brought up as presentations are indicated by a P. Occasionally, a topic, labeled QP, was raised as a question, but because the student did not give a correct answer, the tutor provided the answer. It can be seen that in general the questions occurred at the top levels of the outline and at the initial topics within each level rather than at the later topics. We think this was because the top-level topics and the initial topics within a level were the most important topics, and since the tutor thought the student was likely to know about them, he asked about them rather than presenting them as new material.

As an explicit theory of the interweaving of questioning and presentation, we would argue that when the tutor thinks the student

may know the answer, he introduces the topic as a question rather than presenting the information. The tutor's evaluation of whether the student will know the answer is made using an a priori estimate of the sophistication of the student, which is refined as the dialogue progresses, based on the answers and information the student provides. This evaluation by the tutor of the student's sophistication would be based on something like SCHOLAR's I-tags. As the student answers some questions and fails to answer others, both the answers and failures will have levels of importance with respect to the top-level topic, South America, in the tutor's own data base. It might turn out, for example, that the student can always answer questions where the importance level of the answer is from 0 to 2, sometimes answer when 3 or 4, and never answer when 5 or more. Then it would be a sensible strategy to assume that the student knows information with a level less than 2, to ask questions if the information has a level of 3 or 4, to present information if the level is 5 or 6, and to omit information if the level is 7 or more.

The student's ability to answer will correspond quite well with the tutor's ranking of importance. This is because there is common cultural agreement as to what is important and what is not, and learning typically proceeds from the more important to the less important. There are sometimes exceptions, as when a person happens to have learned much more about one particular concept than about other similar concepts. In that case, the tutor may be forced to revise his ideas, but he starts out with the assumption of

uniformity in the student's knowledge.

We have suggested above that there are four categories of information, the boundaries of which vary depending on the sophistication of the student. These categories are shown in Figure 5, with examples of the kinds of information that would fall into each one for the student in the dialogue of Figure 3. The first category is what the tutor assumes the student knows. The first tutor never asked any of the four students what South America was, but he might have if the student had been, say, a child of seven. The second category is made up of facts that the tutor tries to elicit with appropriate questions, as with the Amazon and the Andes in the dialogue. In the third category are those facts he does not bother to try to elicit with questions, because he does not think the student will know them. Instead, he presents the information to the student, as he did with the Parana River and Aconcagua in the dialogue, because he thinks the student should be able to assimilate these facts. The fourth category consists of those facts that he does not even present, because they would be more than the student could learn. We infer the existence of a fourth category because some facts are presented to more sophisticated students that are not presented to less sophisticated students. The two facts shown in the figure were presented to the second student, who was quite sophisticated in geography, but not to the student in the dialogue in Figure 3. We assume the levels of these categories move higher or lower together with respect to the tutor's own scale of importance depending on his estimate of the student's sophistication.

	Categories of Information	Examples from the Dialogue in Figure 3
1	Information the tutor regards as very important, which he assumes the student knows, and so does not ask about.	South America is a continent. South America is south of North America.
2	Information the tutor regards as important, which he thinks the student may or may not know, and therefore asks about.	The Andes are the major mountain range in South America. The Amazon is a large river in South America.
3	Information the tutor regards as somewhat less important, which he thinks the student probably does not know, and so he presents the information to the student.	The Parana is a large river in South America. The highest mountain in the Andes is Aconcagua.
4	Information the tutor regards as still less important and too much beyond the student's level of sophistication to be worth presenting.	The Paraguay River is a tributary of the Parana River. Manaus is a port half-way up the Amazon River.

Figure 5: The Different Categories of Information that Determine What Questions are Asked and What Information is Presented.

This description of the tutor's evaluation of the student is related to the problem which Norman (1973) refers to as the Empire State Building Problem. As he points out, the answer to "Where is the Empire State Building?" depends on where a person asks the question. In Russia the appropriate answer might be "The United States"; in England it might be "New York City"; and in New York City it might be "On 34th Street". Norman suggests that to answer such a question appropriately, "it is necessary to have a model of the knowledge of the listener". What we are presenting here is a fairly simple mechanism by which a person can adjust the level of his answer (or in our case, his question or presentation) to the level of sophistication of the other person. In terms of the Empire State Building Problem, the strategy would work as follows: People have various pieces of information stored about the location of the Empire State Building, with various levels of importance. By estimating how sophisticated the other person is (Norman suggests that where he is is one criterion for estimation), an appropriate answer can be selected. The appropriate answer would have a low I-tag for a Russian, a higher tag for an Englishman, and a still higher tag for a New Yorker. Of course, it might turn out that a Russian questionner had spent several years in New York, but his reply would quickly reveal his knowledge about New York. Any misestimate of a person's sophistication can be corrected by evaluating the level of the information he provides in his reply.

We mentioned in the first paragraph of this section that the tutorial strategy for interweaving questions and presentation is

quite different from the strategy employed in programmed learning, on which most CAI systems are based. The programmed learning strategy involves presenting small amounts of information and then asking questions about that information. Because this strategy cannot be geared to the prior knowledge of the student, most programs using it start at a fairly low level. For this reason the student often finds himself going over material he already knows, which is boring. But for a less sophisticated person the same material might be too difficult. Even when the material is at about the right level, the student is giving answers based on material recently presented to him. Thus, he often winds up half parroting what he has just read, a mode of recall that Craik (1970) and Madigan and McCabe (1971) have shown leads to little or no long-term retention.

In contrast, the questioning in the tutorial strategy precedes the presentation of material. The questioning determines what the student knows about a particular topic, and then semantically related material is presented which goes a little beyond the level of knowledge the student has shown in answering the questions. Thus the tutor can build onto the knowledge the student already has, without going beyond what the student can assimilate. This is an essential aspect of the way the tutor gears his teaching to the level of the student.

Questioning about Basic Concepts. There was another kind of questioning that occurred in the dialogues, which we did not

anticipate. This was the questioning by the tutor as to whether the student knew the meaning of some basic geographical concepts. Such questions occurred four different times in the dialogue in Figure 3; in particular with respect to the meaning of estuary, delta, savanna, and temperate climate (lines 71, 75-79, 107, and 151). The questions arose when these basic concepts were relevant to the discussion of particular topics, such as the Rio de la Plata. Hence, they could occur at any time and were not tied to the structure of the outline as were the factual questions discussed in the previous section[6].

As with factual information, the level of sophistication of the student seems to determine which basic concepts a tutor asks about. Whereas the tutor asked the student in the first dialogue about all four basic concepts mentioned above, he asked the more sophisticated student in the second dialogue only about the meaning of savanna. This was true even though all of the other concepts (i.e., delta, estuary, and temperate) came up in this second dialogue. In tutoring children, the same tutor probably would have asked about even more basic terms like government or plains.

With respect to basic concepts, the tutors seemed to be working only at the first two categories of the four shown in Figure 5. That is, they either assumed the student knew the basic concept or else they asked the student if he knew what it meant. However, with savanna in Figure 3, the tutor phrased the question in a way that presupposed that the student did not know the concept (line 107,

"Have you heard of that term before?"). In no case did a tutor explain to the student what a basic concept meant without asking the student about it. The latter strategy was, on the other hand, often seen with factual knowledge, as illustrated in the third level of Figure 5. Our guess as to why this strategy did not occur is that the few concepts brought up were fairly basic and the students were all adults. If the first tutor had been teaching a child, he probably would not have bothered to ask the child what a savanna was, but instead would have explained without asking. We would argue that the same mechanism is used with both factual knowledge and basic concepts.

Reviewing by the Tutor. When the first tutor finished his discussion of South America as a whole in Figure 3 (line 115) and had started discussing each individual country, he asked the student a whole series of questions in review (lines 117-131). There are two separable aspects to this kind of reviewing: reiteration and review questions.

By reiteration we refer to repeated passes through the same topics. It is systematic and it has to do with the overall organization of the session. The repeated passes may involve review questions or introduction of new material. By review questions we refer to the tutor's questions about material covered earlier. Review questions sometimes occurred in the framework of systematic reiteration; they also frequently arose when an old topic came up in a new context.

Reiteration often occurred in the dialogues (as in the one shown) when the tutor went over much of the same material discussed under South America, but on a country-by-country basis. In going through a second time, he questioned the student about some of the old material (though tutors sometimes omitted these questions), and at the same time he added related new material, such as that about Tierra del Fuego (line 135) and the Chaco (lines 137-141). He could have discussed such topics in relation to South America as a whole on the first time through, but this would have increased the amount of new information for the student to assimilate on the first pass. By reiterating in this way the tutor can approximate what Norman (1973) refers to as "web teaching".

As Norman describes it, the object of web teaching is to establish a coarse web of interrelated material that is well integrated with previous knowledge. When this is done, new pieces of information can be added by tying them to the original web framework. The process can be repeated over and over, adding more and more detail each time. In addition to helping the student assimilate the new information, web teaching follows an order in which the most important information is taught first and information that is successively less important is taught on later passes. If the teaching is not completed for some reason, the most important information will still have been covered. And on later passes the material reviewed will be the more important information taught on

earlier passes. For these reasons, web teaching is probably the most effective method of teaching.

Review questions about previously discussed material did not occur on a systematic basis in the dialogues. For instance, the tutors did not review most of the questions that the student missed during the session. Nor did review questions occur at a systematic place in the dialogue, such as at a fixed time after a question that was not answered. The place where review questions did occur frequently was when a topic discussed earlier came up in another context. This usually happened during a second pass through material discussed earlier, but not always. In one case, for example, a region extending into two countries was mentioned first in discussing one and then a review question was posed later in discussing the other.

The example in Figure 3 is unusual in that there was a whole series of review questions, whereas more commonly there were only one or two review questions at a time. In the series of six review questions the student had answered three correctly earlier, and these three could well have been omitted. This tutor did not ask many review questions in any of the dialogues, and he clearly was not systematically reviewing information he had told the student earlier.

The second tutor utilized review questions more frequently and seemed to come back more often to questions that the student could not answer earlier. But he too only re-asked a question when the

topic came up in another context. The percentage of his review questions which the student had not answered before (eleven of thirteen in the dialogue counted) was higher than for the first tutor, but he did not exhibit perfect memory either. The other two tutors asked very few review questions. None of the tutors followed what is probably the optimal strategy in asking review questions, that is, reviewing all those topics where the information has been provided by the tutor, and not reviewing those topics where the information has been provided by the student.

Providing Hints. There were a number of places in the dialogues where a tutor provided a hint to a student because he thought the student might know the answer. Here again is a strategy designed to individualize the teaching to the student's prior knowledge. One example of the use of hints appears near the beginning of the dialogue shown in Figure 3 (line 15). The tutor tried to elicit Colombia by telling the student that it was connected to Panama. In Figure 6, we show examples of hints from other dialogues. The hints are marked in parentheses; there are two in the first excerpt and one in the second. In the first case, the second hint worked. In the second case, the tutor gave up after a while and provided the correct answer. In both examples, the student had previously read some material on South American geography, and the discussion was a review. Hints seemed to occur most frequently in a review session, though the fourth tutor often used hints in a non-review condition.

In the two examples shown, the tutors provided hints after the student had said "I don't know" or "I can't remember". In each case the tutor must have gauged the level of sophistication of the student to be above the level of difficulty of the answer, or else he would not have pursued the correct answer so assiduously. The fact that the student had already named less well-known countries in the first case, such as Colombia and Paraguay, and a much lesser city, Santos, in the second case, would indicate to the tutors that in each case the student should know the answer.

Though we did not find any examples in the dialogues, there is another situation where hints might profitably be provided. When a subject failed to answer or to say "I don't know", tutors usually waited about five seconds before providing the correct answer. It would be a natural strategy here to give hints rather than the correct answer, because the student's failure to answer one way or another probably means that he thinks he knows, but cannot find the answer.

The function of hints may be twofold: (1) It may help the student to remember a fact if he comes up with the answer himself, rather than hearing it from the tutor. (2) Hints may act to relate the facts given as hints to the name being sought as an answer, and to help the student to remember a whole bundle of interrelated information.

1. T1. And (what is the name of) this country. (points to Venezuela)
2. S. I don't know.
3. T1. This country has Lake Maracaibo in it. (hint) Does that help you? This is Lake Maracaibo. (points)
4. S. No.
5. T1. It is rich in oil deposits. (hint)
6. S. Venezuela.

---

1. T2. Do you know what city this is? (points to Sao Paulo)
2. S. I think it's Santos.
3. T2. No. Which one is Santos? (error correction)
4. S. No. The other one is Santos. Right here. (points to Santos)
5. T2. Yeah. The port city.
6. S. The port city. Right. And I can't remember what that is. (points to Sao Paulo)
7. T2. Now this is the second largest city in South America. (hint)
8. S. Buenos Aires.. No, that's in Argentina.
9. T2. No, Buenos Aires is the largest and that's down in Argentina. (error correction)
10. S. The second largest. And it's the former capital of Brazil.
11. T2. No, this is the former capital of Brazil. (points to Rio) Do you know what city this is? (error correction)
12. S. Yes, Rio de Janeiro.
13. T2. That's Rio de Janeiro and this (points) is Sao Paulo, and it's almost twice as big as Rio.

Figure 6: Fragments of Dialogues Where the First Two Tutors Used Hints.

The Tutors' Response to Errors. One of the two aspects of tutorial strategies that we initially planned to investigate was how the tutor dealt with errors. Before we looked at the dialogues, we thought that the tutor might respond to an error by questioning the student to determine the underlying misconceptions that produced the error. None of the tutors responded in this way, perhaps because the errors were all obvious ones. Interestingly enough, each of the three tutors who were faced with errors by students responded to the errors in a somewhat different pattern. (The student in the dialogue with the fourth tutor never committed herself enough to make errors.)

The error correction strategy that the first tutor used most often was to point out which of two things was which, and then to provide distinguishing characteristics. This can be seen in two cases during the naming of countries at the beginning of the dialogue shown in Figure 3 (lines 13-15 and 21-27). In the first example, the tutor corrected the student's confusion between Colombia and Ecuador (line 15). For Ecuador he pointed out that it was the smaller country, and for Colombia that it connected with Panama. (The latter fact was offered as a hint by the tutor to see if he could elicit the correct answer.) In the second example (lines 21-27) the student confused Paraguay and Uruguay, and again the tutor provided some properties of Uruguay that could be used to distinguish Uruguay from Paraguay. Providing one or two distinguishing characteristics between the correct answer and the wrong answer was the essence of the first tutor's strategy for

correcting errors.

Like the first tutor, the second tutor usually provided a distinguishing property to undo a student's confusion, but in addition, he frequently asked a question about the student's wrong answer as well. Three examples of this tutor's correction of errors are shown in the fragment of the dialogue at the bottom half of Figure 6. In the first example, the student identified Sao Paulo as Santos in line 2, but the tutor did not tell him the correct name until line 13. Before providing the correct answer, the tutor asked where Santos was (line 3), and then mentioned the distinguishing property that Santos was the port city (line 5). The second example occurs in lines 8 and 9, where the student mentioned Buenos Aires and then corrected himself. The student's self-correction precluded a question about Buenos Aires ("No. Where is Buenos Aires?"), but the tutor still pointed out that Buenos Aires was the largest city, not the second largest city. In the third example (lines 10-13), the student suggested that the city in question, Sao Paulo, was the former capital, while the former capital was in fact was Rio. Here again the tutor formulated a question about Rio and also provided the distinguishing property that Sao Paulo was much larger than Rio.

The strategy of responding with a question about the wrong answer was probably used to help the student remember the distinction. Presumably, by generating the correct answer himself, the student would remember it better. In contrast, the first tutor probably would have responded to the first error (line 2) by telling

the student where Santos was and mentioning its most salient property (i.e., that Santos was the port city for the city in question). The second tutor did mention the fact that Santos was the port city, but only after the student answered the question about the location of Santos. The second tutor's strategy for correcting errors consisted of first asking a question about the wrong answer (frequently a "Where is it?" question) and then providing a distinguishing property between the correct and the wrong answer.

The third tutor encountered only six errors in his one dialogue, and no obvious strategy emerged other than pointing out the error. In five of the six cases, he followed just the minimal strategy of giving the correct answer. In the sixth case he too used the strategy of providing a distinguishing property in a case where the distinction (in particular between Uruguay and Paraguay) was difficult to remember.

By providing distinguishing characteristics, the tutors were directly combatting the students' confusions. The strategy is a good one, because the essence of discrimination learning is finding properties that may be used to distinguish the two things in the future.

4. SYNTHESIS OF TUTORIAL DIALOGUES

Based on our analysis, we have implemented tutorial strategies similar to those described above, in a Tutorial Mode of SCHOLAR. In

particular we have modelled the tutor's strategies for: (1) topic selection, (2) the interweaving of questioning and presentation, (3) reviewing, and (4) error correction. We haven't always simulated the analysis exactly; this was not because it was impossible, but rather because the structure of SCHOLAR made it easier to use different means to accomplish the same basic objectives. In this section, then, we will try to emphasize in what ways we have captured these strategies in SCHOLAR, and in what ways they have been altered in the implementation.

Figure 7 illustrates the Tutorial Mode in SCHOLAR with an annotated transcript of a dialogue between a high-school student and SCHOLAR. The transcript has been retyped to save space, but it is verbatim. The high-school student had no particular familiarity with computers or with SCHOLAR. She was given examples of the types of questions she could ask, but otherwise she was essentially naive. This is a typical dialogue with a well-motivated student. One of the more impressive aspects of SCHOLAR is that it can carry on a reasonable conversation with someone who knows practically nothing about it.

4.1 Comparison of Computer and Human Dialogues

The comparison of the machine dialogue in Figure 7 and the human dialogue in Figure 3 should be made in terms of the aspects we have tried to model in SCHOLAR: topic selection, the interweaving of questioning and presentation, reviewing, and error correction. We

will discuss each of these aspects below.

The comparison should not be made in terms of wording or naturalness. Though these are relevant considerations for judging SCHOLAR, we should emphasize that it will be a long time before any machines are as flexible as humans in the use of language. On the other hand, in spite of this limitation, generative CAI systems like SCHOLAR may be useful already for tutoring such subjects as geography or electronics (Brown, Burton, and Bell, 1974).

Topic Selection. Tutorial Mode of SCHOLAR (hereafter called SCHOLAR) selects topics according to the five rules given in Section 3.2, as modified to handle embedded attributes (see footnote 4). Like the tutor, SCHOLAR takes up topics in the order of their importance, and it discusses each subtopic that comes up during the dialogue for a little while before going on to more important topics. But there is more variability from dialogue to dialogue in the topics that the human tutor brings up. Or, put in the opposite light, SCHOLAR is more thorough about covering all the major subtopics under any topic.

In selecting subtopics under the main topic, SCHOLAR always picks the most important top-level attribute that has not yet been discussed. Then it goes down into the information under that top-level attribute, always picking the most important subattribute until it finds a value. Near the beginning of the dialogue shown in Figure 7, SCHOLAR went from the main topic Chile, to the top-level attribute Location, down into Bordering-Countries, further down into

WHAT IS CHILE?

COUNTRY IN SOUTH AMERICA

VERY GOOD.

This is the most important property of Chile. (A tutor would suppress it on the assumption that the student knows the answer.) In evaluating the answer, it looks for "country", ignores the rest.

CORRECT OR INCORRECT?
CHILE IS IN NORTHERN SOUTH AMERICA.

INCORRECT

YOU ARE CORRECT.

CHILE IS IN SOUTHWESTERN SOUTH AMERICA.

Location is the next most important attribute under Chile.

The student answers correctly, so SCHOLAR will go deeper into Location.

THE EASTERN BORDERING COUNTRY OF CHILE IS ---.

I DON'T KNOW

THE ANSWER IS: ARGENTINA.

This is the most important Bordering-Country of Chile.

THE BOUNDARY OF THE EASTERN BORDERING COUNTRIES OF CHILE IS THE ANDES.

THE NORTHERN BORDERING COUNTRY OF CHILE IS PERU.

THE NORTHEASTERN BORDERING COUNTRY IS BOLIVIA.

Since the student has missed the question, SCHOLAR presents some related information about Bordering-Countries.

WHAT IS THE CAPITAL OF CHILE?

WHAT ARE SOME CITIES IN CHILE

THE PRINCIPAL CITIES OF CHILE ARE SANTIAGO, VALPARAISO, CONCEPCION, ANTOFAGASTA, ARICA, VINA DEL MAR, PUNTA ARENAS, AND IQUIQUE.

SCHOLAR would have gone on to ask about Coasts under Location, but it has run out of time for Location, so it goes to the next attribute, Capital.

The student gets some help, though without directly asking SCHOLAR's question.

Figure 7a: An Annotated Transcript of a Dialogue between Tutorial Mode of SCHOLAR and a High-school Student. (The student's inputs are between asterisks.)

NOW YOU CAN EITHER ANSWER MY QUESTION
OR ASK A NEW ONE.

SANTIAGO

VERY GOOD.

Because the student is right, SCHOLAR goes into the data base entry for Santiago to ask more questions. (To go into an entry, the entry must have the main topic as a Superpart, i.e., it must be part of Chile.)

IS THE FOLLOWING TRUE OR FALSE?
SANTIAGO IS IN CENTRAL CHILE.

TRUE

THAT'S FINE.

SCHOLAR has run out of the time allotted to Capital and so goes on to Regions. Time is allocated on the basis of the number of top-level attributes and their I-tag weights.

THE REGIONS OF CHILE ARE ---.

I DON'T KNOW

THE ANSWER IS: THE CENTRAL VALLEY, THE ATACAMA DESERT, AND THE SOUTHERN ANDES.

Since the student does not know any Regions, SCHOLAR gives the correct answer, plus some information about two of the Regions.

THE CENTRAL VALLEY IS IN CENTRAL CHILE.
IT IS ON THE COAST OF THE PACIFIC.

THE ATACAMA DESERT IS IN CHILE AND PERU.
IT IS BETWEEN THE PACIFIC AND THE ANDES MOUNTAINS.

Language is under the top-level attribute Population. SCHOLAR goes on to Population not because of time but because the student doesn't know any Regions.

WHAT IS THE LANGUAGE OF CHILE?

SPANISH

YOU ARE CORRECT.

That's all SCHOLAR knows about Language, so it goes on to Religion, which is at the same level as Language under Population.
The rest of the information about Religion is above the I-tag level for the first pass, so no related information is printed.

WHAT IS THE PRINCIPAL RELIGION OF CHILE?

I DON'T KNOW

THE ANSWER IS: CATHOLICISM.

Figure 7b: An Annotated Transcript of a Dialogue between Tutorial Mode of SCHOLAR and a High-school Student.

IS IT TRUE THAT THE MINERALS OF CHILE ARE VERY IMPORTANT?

WHAT ARE SOME OF THE MINERALS?

THAT WAS MY QUESTION.

THE MINERALS ARE VERY IMPORTANT.
THE PRINCIPAL MINERALS ARE COPPER, IRON, SILVER, COAL, NITRATES, AND IODINE.

The time allotted to Population is used up, so SCHOLAR goes on to Minerals under Products.

SCHOLAR understands this as asking about the Minerals, so it points out that it has just asked about them.

IS IT CORRECT TO SAY THAT THE MOUNTAINS OF CHILE ARE NUMEROUS?

TRUE

VERY GOOD.

Mountains are stored under Topography. SCHOLAR goes on to Topography because the time for Products is used up. The higher I-tag on Products means less time is allocated.

WHAT ARE THE PRINCIPAL MOUNTAINS OF CHILE?

I DON'T KNOW

THE ANSWER IS: THE ANDES.

THE HIGHEST MOUNTAIN OF CHILE IS THE ACONCAGUA.

Since the student has answered the first question about Mountains, SCHOLAR goes deeper to ask the name. When the student misses that, SCHOLAR gives the correct answer plus a related piece of information.

IS IT TRUE THAT THE GOVERNMENT OF CHILE IS UNSTABLE?

YES

THAT'S FINE.

Because of time, SCHOLAR skips over Rivers, Lakes, and Soil under Topography and goes on to Government.

IS IT CORRECT TO SAY THAT THE GOVERNMENT OF CHILE IS MILITARY?

YES

YOU ARE CORRECT.

Figure 7c: An Annotated Transcript of a Dialogue between Tutorial Mode of SCHOLAR and a High-school Student.

RIGHT OR WRONG?
THE GOVERNMENT OF CHILE IS DICTATORIAL.

RIGHT

THAT'S FINE.

Topics with I-tags less than the prespecified level have now been covered, so this ends the first pass.

THE EASTERN BORDERING COUNTRY OF CHILE IS ---.

I DON'T KNOW

On the second pass, SCHOLAR skips to the first piece of information the student missed earlier.

THE ANSWER IS: ARGENTINA.

THE BOUNDARY OF THE EASTERN BORDERING COUNTRIES OF CHILE IS THE ANDES.

THE NORTHERN BORDERING COUNTRY OF CHILE IS PERU.

THE NORTHEASTERN BORDERING COUNTRY IS BOLIVIA.

Since the student misses it again, the same related information as before prints out. SCHOLAR knows that this information was presented earlier, and we may revise SCHOLAR so that it asks about the related information here instead of presenting it again.

WHAT ARE THE REGIONS OF CHILE?

CENTRAL VALLEY ATACAMA DESSERT AND THE ANDES

"CENTRAL VALLEY" AND "ATACAMA DESERT" ARE FINE.
YOU COULD ALSO MENTION THE SOUTHERN ANDES.
I DON'T THINK "DESSERT" IS CORRECT.

THE ATACAMA DESERT IS IN CHILE AND PERU.
IT IS BETWEEN THE PACIFIC AND THE ANDES MOUNTAINS.

THE SOUTHERN ANDES ARE IN SOUTHWESTERN SOUTH AMERICA.
THEY ARE IN ARGENTINA AND CHILE.

The student missed this the first time. A new answer evaluation module will soon point out that the Andes is right but more specifically it is the Southern Andes.

SCHOLAR recognizes Atacama as the Atacama Desert and does not understand Dessert.

Why it gives related information about the Atacama when the student has gotten it right is a mystery.

Figure 7d: An Annotated Transcript of a Dialogue between Tutorial Mode of SCHOLAR and a High-school Student.

IS IT TRUE THAT THE CENTRAL VALLEY IS IN CENTRAL CHILE?

YES

VERY GOOD.

Because the student has named the Central Valley, SCHOLAR goes into the entry for Central Valley to ask questions about it. Like the tutors, SCHOLAR will go off and discuss what the student answers.

CORRECT OR INCORRECT?
THE CENTRAL VALLEY IS ON THE COAST OF THE PACIFIC.

INCORRECT

WRONG.

THE CENTRAL VALLEY IS ON THE COAST OF THE PACIFIC.

This is all the information it has about the Location of the Central Valley, so no related information is given.

IS IT CORRECT TO SAY THAT THE CLIMATE OF THE CENTRAL VALLEY IS TEMPERATE?

WHAT IS TEMPERATE?

TEMPERATE CAN BE CONSIDERED A PART OF A CLIMATE.

YOU MAY ASK MORE QUESTIONS IF YOU WANT TO.

This is a failure on our part to get around to defining Temperate properly in the data base. It is for this kind of question that the interactive ability of SCHOLAR should be most useful.

IS TEMPERATE VERY WARM CLIMATE?

I DON'T KNOW WHAT IT MEANS FOR VERY TO HAVE TEMPERATE.

FEEL FREE TO ANSWER THE QUESTION YOU DIDN'T ANSWER BEFORE.

This question is beyond SCHOLAR's capability to analyze, but it seems to have tried anyway. More usually it would say "I don't understand your question."

I DON'T KNOW

THE ANSWER IS: TRUE.

THE CLIMATE OF THE CENTRAL VALLEY IS MEDITERRANEAN.

This is a related fact about the Climate.

Figure 7e: An Annotated Transcript of a Dialogue between Tutorial Mode of SCHOLAR and a High-school Student.

WHAT ARE THE PRINCIPAL AGRICULTURAL PRODUCTS OF THE CENTRAL VALLEY?

I DON'T KNOW

THE ANSWER IS: WHEAT, GRAPES, AND WINE.

That is all it knows about Agricultural Products.

IS IT TRUE THAT THE MINERALS OF THE ATACAMA DESERT ARE IMPORTANT?

NO

When it pops out of Central Valley, it goes into the Atacama Desert, since the student also named it as a Region.

YOUR ANSWER IS INCORRECT.

THE MINERALS OF THE ATACAMA DESERT ARE IMPORTANT.

THE PRINCIPAL MINERALS OF THE ATACAMA DESERT ARE COPPER, IRON, NITRATES, SILVER, SULFUR, AND ZINC.

Since the student misses the top-level question about Minerals, SCHOLAR tells him the Minerals there.

(The dialogue continues on through the second pass and part of a third.)

Figure 7f: An Annotated Transcript of a Dialogue between Tutorial Mode of SCHOLAR and a High-school Student.

Eastern, until it found the value Argentina. It then formulated a question about the eastern bordering country of Chile. In going from Bordering-Countries into Eastern, it chose Eastern rather than Northern because of a difference in I-tag values. At each embedded level of the data base, SCHOLAR is always choosing the most important branch, as determined by I-tags.

Of the two ways we suggested that a tutor is influenced by context, SCHOLAR incorporates one and not the other. The major effect of context in the human dialogues was that each subject that came up in the discussion was potentially the next topic. This occurs in SCHOLAR just as the five rules provide. Perhaps the best example in Figure 7 is in the two cases where SCHOLAR asked about the regions of Chile. The first time, the student couldn't name any regions, so SCHOLAR told her their names, plus a few relevant facts about two of them. Then it went on to other topics. The second time, near the end of the dialogue, the student correctly named the Central Valley and the Atacama Desert, so SCHOLAR began asking a series of questions about each of these (after it had provided a little information about the Southern Andes, which she had missed). Thus, the answers that the student happens to give determine the topics that SCHOLAR will take up. This parallels the situation in the dialogue in Figure 3, where the student mentioned Cape Horn in answer to the question about Geographical-Features, and Cape Horn became the topic for several minutes of discussion.

The second way that context appeared to influence the human

dialogues was in the selection of a new topic after popping up from the previous topic. This does not occur at all in Tutorial Mode, though something like it occurs in the original Mixed-Initiative Mode of SCHOLAR (Carbonell, 1970a, 1970b). In Mixed-Initiative Mode, if the student asks a question about some topic, the random selection of topics is biased toward selecting a new question on the topic which the student brought up. A more elaborate scheme might change I-tag values depending on context (see Carbonell and Collins, 1973). But Tutorial Mode makes its selections on the basis of fixed I-tag values, and this accounts for its inflexibility or thoroughness (whichever you prefer to call it) in selecting topics.

The decision as to when to go down deeper as against when to pop up to a new topic is determined by several conditions. These form the criterion referred to in Rule 2 of the five rules. The tutors appeared to take both importance and available time into consideration in deciding when to pop up. There is no way to tell from the dialogues how they combined the two, so we have adopted a complicated trade-off, with no theoretical implications. Briefly the scheme is this: Time is allocated to each of the top-level attributes in proportion to its importance. Similarly, in allocating the time for subtopics under an attribute, such as the Central Valley and the Atacama Desert under Regions of Chile, time is allocated proportionately. When the time allocation at any level is used up, there is a pop-up and the next most important concept at that level is selected. There is also a cutoff level in terms of importance (currently set at an I-tag level of five). This can be

adjusted depending on the time available; to go deeper if SCHOLAR is running ahead of time, and to go less deep if it is running behind. In the dialogue in Figure 7, it was set at a fixed level. Pop-ups also occur when the student misses a question, but that is discussed in the next section. If a pop-up occurs because of a missed question or an I-tag cutoff, any extra time left is added to the allocation for the next topic.

Interweaving of Questioning and Presentation. Like the human tutor, SCHOLAR starts off questioning the student. Then it presents some new information related to what he already knows. The object is to tie the new material into the old, and to give the student as much information as he can assimilate at one time.

This is achieved in SCHOLAR by a somewhat different strategy from the one described for tutors. SCHOLAR does not form a model of the student, other than to build an event memory of who said what during the dialogue. It does not estimate the student's sophistication, nor operate with the four categories described for tutors. Instead it starts out asking questions, going down to deeper and deeper embedded levels until either the student cannot answer correctly, or one of the criteria for popping up is met. When the student cannot answer a question, SCHOLAR presents two or three related facts that are embedded within that attribute, and then backs up to the next most important attribute at the same level as the question missed.

There are three major ways that this produces a different kind of dialogue from the human dialogues. First, SCHOLAR will ask questions about all its important information (e.g., What is Chile?), whereas a tutor would skip over any information he assumed the student would know.

Second, the tutors sometimes presented information when they brought up new topics, as did the first tutor when he discussed the Uruguay River and the Orinoco River. SCHOLAR, on the other hand, will always introduce a new topic with a question. Tutors therefore tended to talk for longer periods of time without asking any questions than does SCHOLAR.

Third, SCHOLAR pops up more often than the tutors, because of the time cutoff. Thus the pace of the conversation was more leisurely in the human dialogues, because the tutors would go into most topics in more detail. On the other hand, SCHOLAR covers a greater variety of topics, because it is always moving on to new topics. The one place in Figure 7 where SCHOLAR did not hurry on to new topics was in the discussion, near the end, of the Central Valley and the Atacama Desert. This was because there was time left over from previous topics. It would be easy to relax the time constraint, but it is not obvious which is the better strategy.

Reviewing. We distinguished two aspects of reviewing in our analysis of dialogues; they were reiteration and review questioning. In Tutorial Mode we modelled the tutor's behavior in both respects.

Reiteration is essential to the strategy of web teaching, which we think may be the most effective method of teaching. Therefore, we set up Tutorial Mode to allocate time so as to provide for at least two passes through the material. If there is time left at the end, it goes through more passes. Reiteration by the human tutors usually occurred by having a first pass on South America, and a second pass on each country. In SCHOLAR, we are teaching about each country individually, and so each pass covers the same material.

We could have set the I-tag cutoff higher on each subsequent pass in order to go into the material in more depth. However, this effect occurs anyway, for two reasons. Time is saved on the second pass by skipping over material that has already been answered correctly, and so SCHOLAR has more time to use up. Also, the students miss fewer questions on later passes. For both these reasons, there are fewer pop-ups on the later passes, and the material is covered in more depth.

With regard to review questions, SCHOLAR follows the optimal strategy we described in the analysis. That is, it asks about information the student missed earlier, and it skips over information the student knew earlier. But when the student misses the same question twice, SCHOLAR prints out the same few related facts each time. A better strategy might be to ask about these facts the second time through, rather than to present them again. However, if the student answers the question correctly the second time, SCHOLAR will in fact ask about the related facts that were

presented earlier.

Like the tutors, SCHOLAR sometimes returns to a topic discussed earlier, at times other than on a second or third pass. This happens when the same topic comes up in a different context. For example, the Pampas might come up first under Regions of Argentina, and later as a Plain under Topography of Argentina. Just as in a second pass, SCHOLAR will repeat those questions that the student missed the first time, and will skip over those questions that he answered the first time. Hence, SCHOLAR also asks review questions independent of reiteration.

Error Correction. Based on the analysis of the dialogues, we added an error-correction subroutine to SCHOLAR to provide some types of distinguishing properties when the student makes a confusion. Although no examples occurred in Figure 7, we showed two examples in Figure 2, where SCHOLAR pointed out differences between the student's answer and the correct answer. In one case it pointed out a difference with regard to the Superordinate, and in the other case with regard to the Superpart. In particular, when the student named Buenos Aires and Titicaca as cities in Brazil, the error-correction subroutine pointed out that Titicaca was a lake, not a city, and Buenos Aires was in Argentina, not in Brazil. The capability to find these two types of distinctions was added to SCHOLAR as a result of the analysis of the tutors' strategy.

We have recently developed a much more sophisticated answer-evaluation and error-correction module[7] that says "No" only

if it can find some basis for distinguishing the student's answer from the correct answer. It then tells the student the basis for distinguishing the two in order to correct the student's confusion. Several sample outputs are shown in Figure 8.

The first example illustrates the output when the two things confused are stored under the same top-level attribute in the data base. In this case, SCHOLAR first checked that the Chaco and the Pampas were not the same. Then, since the Chaco, as well as the Pampas, was stored as a Plain under Argentina, SCHOLAR printed out the information stored under Argentina about the Chaco, as the most relevant distinction.

The second example also concerns the Chaco, and illustrates the fact that true/false questions by the students require the same processing (but usually different comments). Here the mismatch with the Chaco occurred on the basis of the countries it was part of, and this distinction was pointed out. The distinction here is different from the one in the first example, though both examples involve the location of the Chaco.

The third example illustrates how the module handles a slight mismatch, like the one that occurred in the dialogue of Figure 7. Southern Andes is stored as a region of Chile, and it has the student's value, Andes, as a Superpart. In the search for a distinction between the two, the module found this relation, gave a weak yes, and indicated the more specific value that was stored.

```
 THE EASTERN AND CENTRAL PLAIN OF ARGENTINA IS ---.

*CHACO*

 NO.
 THE NORTHERN PLAIN OF ARGENTINA IS THE CHACO.  THE EASTERN
AND CENTRAL PLAIN OF ARGENTINA IS THE PAMPAS.


*IS THE CHACO A REGION IN CHILE?*
 NO.
 THE CHACO IS IN ARGENTINA, PARAGUAY, AND BOLIVIA, NOT IN
CHILE.  THE REGIONS OF CHILE ARE THE ATACAMA DESERT,
THE CENTRAL VALLEY, AND THE SOUTHERN ANDES.


*ARE THE ANDES A REGION IN CHILE?*
 WELL, YOU'RE ALMOST RIGHT..
 MORE SPECIFICALLY, A REGION OF CHILE IS THE SOUTHERN ANDES.
```

Figure 8: Examples of Output from New Answer Evaluation Module

There are a large number of possible outcomes from this module, but these will be described in detail elsewhere (Collins, Warnock, and Aiello, in preparation). The few examples shown here illustrate that the processing involved in deciding that a student's answer is wrong leads to the most relevant distinction. Being told about that distinction can help the student to distinguish the two things in the future.

4.2 Use of SCHOLAR to Evaluate Teaching Methods

There is a useful way that a program like SCHOLAR can be used in research on teaching strategies. It is possible to run SCHOLAR in different modes when it is teaching students. For example, SCHOLAR can operate in a mode where it first presents a block of information and then asks questions about it (Block-Test Mode). Therefore, we can test whether students learn more with this method of teaching or with the tutorial method, when the information covered is exactly the same in both. To measure any such effect, we compare students' improvement in scores on a test administered before and after sessions with these two modes in SCHOLAR (see Collins, Passafiume, Gould, and Carbonell, 1973)[8].

There are many different teaching methods that could be compared in this way. For example, when a facility for teaching geography with maps is completed, we can compare how well students learn the same material with maps and without maps. The possibility of trying out single modifications in teaching strategy to see their

effects on students' learning is unique. Human teachers, of course, can make such modifications in their own teaching strategies, but there is no way to control all the other factors that might vary as they changed strategy. However, in any specified version, SCHOLAR is a fixed system, and so an unbiased comparison can be made using any number of subjects. After testing out single modifications one at a time, it is possible to start combining those factors which show positive effects on students' learning, and to test them out in combination. In this way we can begin to accumulate systematic knowledge about teaching methods.

5. DISCUSSION

Because this is in part a paper on methodology, we would like to conclude it with some comments about the method of dialogue analysis and the general approach of mixing computer synthesis with psychological analysis.

5.1 Comments on Dialogue Analysis as a Method

We have not attempted to make a complete analysis of the dialogues we collected. One could derive much more information from the dialogues and at the same time treat them more systematically than we have here. But we are frankly interested in the dialogues from an applied point of view, and a finer-grain analysis could cloud the important aspects behind a myriad of detail. Therefore, we have only looked at those aspects that seemed particularly

relevant to the way the tutor relates his teaching to the individual student.

There are many other questions that could be investigated using dialogue analysis. In the limited setting of tutorial dialogues one could ask the following questions, for example: (1) Based on the percentage of correct answers on a post-test, what variables of the dialogue affect whether or not the student remembers what he was told? (2) What types of questions (e.g., true/false, multiple choice, "WH" questions, etc.) do tutors use and what types do students use? (3) If given instructions to do so, can the tutor keep reviewing all the facts that were covered in, say, the session's first half hour, until the student recalls each fact at least once? (4) What differences in strategies do tutors use to teach different kinds of knowledge such as factual knowledge, functional knowledge, procedural knowledge, and visual or pictorial knowledge? These examples illustrate some of the variety of ways dialogues can be analyzed; question 1 involves relating dialogues to other data, question 2 involves tallying different cases that meet given conditions, and question 3 involves putting boundary conditions on a participant's behavior. Answers to any of these questions would be helpful to us in building a computer system to tutor students.

In relation to question 4 above, we have studied the tutoring of procedural knowledge (Collins, Passafiume, Gould, and Carbonell, 1973); in particular we studied how to perform various tasks with a text-editing computer system. In this analysis we found several

strategies that were particularly useful for tutoring procedural knowledge, that we had not found in the tutorial dialogues about geography. This merely emphasizes the incompleteness of the analysis presented here. Using other tutors and other topics would surely reveal more such strategies for individualizing instruction.

Another way we have used dialogue analysis was to study the use of inference by tutors and students. We investigated this directly in one session by having the student with the most knowledge about South American geography ask difficult questions of the second tutor. The difficulty of the questions often forced the tutor to make inferences on the basis of his incomplete knowledge. This analysis is reported elsewhere (Carbonell and Collins, 1973; Collins, Warnock, and Aiello, in preparation).

Beyond tutorial dialogues there are many other kinds of conversations which might be explored with dialogue analysis. For example, it would be very useful, in constructing information retrieval systems, to analyze conversations where a person tries to find out from an expert what references exist on a given topic in his field. By looking at the ways that the two people resolve the issue of just what kind of information the person is after, we would have a much better idea of how to organize information retrieval systems and what kinds of interaction would be useful. Another question one could investigate, taken from psychology, is the problem of reference (see Olson, 1970). To study this problem, one could set up dialogues to see the different ways that people refer

to a given object in conversation with people from different backgrounds and in different situations. Or, given the concern with ambiguity among linguists (e.g., Chomsky, 1965), it would be profitable to study in what conversational situations ambiguities appear, and how they are then resolved.

A great many questions about language are in fact questions about dialogues, because language by its very nature involves communication between people. And yet, most research on language, whether in psychology, linguistics, or artificial intelligence, does not even consider the possibility of looking at actual human conversation. An important exception is in the field of language acquisition, particularly the innovative work of Brown (1970, 1973). But traditional laboratory methods in experimental psychology greatly restrict the kinds of questions that can be asked about language. In particular, nothing of an interactive nature can be studied with the present laboratory methods, even though interaction is what language is about. The methods themselves force psychologists to pay attention to what is least relevant about human use of language.

5.2 The Analytic-synthetic Approach[9]

The underlying philosophy in this paper is that the most useful way to analyze how people perform a given task is in synthetic terms, that is, in terms of how that performance could be built into some kind of machine. At the same time, the most productive way to

try to synthesize a machine to perform a task that humans now perform, requires systematic analysis of how people perform the task.

Suppose, for example, that one wanted to formulate a theory of how to build houses. Following a purely analytic approach, one might collect data from watching people build houses, such as the mean and standard deviation of the number of bricks that make up walls, or the order in which the bricks are put in the walls. These variables have something to do with building houses, but not very much. On the other hand, to follow a synthetic approach, you could try to build a house yourself, working out the problems either in advance or as you go along. You might get somewhere this way, but houses are fairly complicated, so you would have a lot of problems and probably not much house. What we are advocating is that it is better to watch people building houses for a while, then run home and try out what you think you saw, and when that falls down, run back and see what you did wrong, etc. You can object that it would be simpler just to ask one of the workers how to build a house. But when it comes to problems in science, there is no one to ask.

As this example might suggest, the reason that the synthetic approach to analysis pays off is that it forces one to pay attention to the relevant variables. The reason that the analytic approach to synthesis pays off is that it helps avoid a lot of mistakes. The combined approach might be called the teeter-totter theory of scientific method.

References

Brown, J.S., Burton, R.R., and Bell, A. An intelligent CAI system that reasons and understands. Cambridge, Massachusetts: Bolt Beranek and Newman, Inc., BBN Report No. 2790, 1974.

Brown, R. Psycholinguistics. New York: Free Press, 1970.

Brown, R. A first language: The early stages. Cambridge, Massachusetts: Harvard University Press, 1973.

Carbonell, J.R. Mixed-intiative man-computer instructional dialogues. Ph.D. dissertation. M.I.T., 1970a. Also BBN Report No. 1971.

Carbonell, J.R. AI in CAI: An artificial intelligence approach to computer-aided instruction. IEEE Transactions on Man-Machine Systems, 1970b, MMS-11, 190-202.

Carbonell, J.R. Artificial intelligence and large interactive man-computer systems. Proceedings of the 1971 IEEE Systems, Man, and Cybernetics Group Annual Symposium, Anaheim, California, 1971.

Carbonell, J.R., and Collins, A.M. Natural semantics in artificial intelligence. Proceedings of Third International Joint Conference on Artificial Intelligence, Stanford, California, August 1973, 344-351.

Chomsky, N. Aspects of the theory of syntax. Cambridge, Massachusetts: M.I.T. Press, 1965.

Collins, A.M., Passafiume, J.J., Gould, L., and Carbonell, J.G. Improving interactive capabilities in computer-assisted instruction. Cambridge, Massachusetts: Bolt Beranek and Newman, BBN Report No. 2631, 1973.

Collins, A.M., and Quillian, M.R. How to make a language user. In E. Tulving and W. Donaldson (Eds.), Organization of memory. New York: Academic Press, 1972.

Collins, A.M., Warnock, E.H., and Aiello, N. Reasoning from incomplete knowledge. To appear in D.G. Bobrow and A.M. Collins (Eds.), Knowledge, understanding and dialogue.

Craik, F.I.M. The fate of primary memory items in free recall. Journal of Verbal Learning and Verbal Behavior, 1970, 6, 143-148.

Madigan, S.A., and McCabe, L. Perfect recall and total forgetting: A problem for models of short-term memory. Journal of Verbal Learning and Verbal Behavior, February 1971, 10, 1, 101-106.

Newell, A., and Simon, H.A. Human problem solving. Englewood Cliffs, New Jersey: Prentice-Hall, 1972.

Norman, D.A. Memory, knowledge, and the answering of questions. In R.L. Solso (Ed.), Contemporary issues in cognitive psychology: The Loyola symposium. New York: Halsted Press, 1973.

Olson, D.R. Language and thought: Aspects of a cognitive theory of semantics, Psychological Review, 1970, 77, 257-273.

Quillian, M.R. Semantic memory. In M. Minsky (Ed.), Semantic information processing. Cambridge, Massachusetts: M.I.T. Press, 1968.

Turing, A.M. Computing machinery and intelligence. Mind. 1950, 59, 433-460. Reprinted in E.A. Feigenbaum and J. Feldman (Eds.), Computers and thought. New York: McGraw-Hill, 1963. 1970, 13, 437-445.

Footnotes

1. This research was sponsored by the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-71-C-0228, Contract Authority Identification Number, NR No. 154-330. Reproduction in whole or in part is permitted for any purpose of the United States Government.

We would like to thank Richard W. Pew, Daniel N. Kalikow and Mario Grignetti for their tutoring, Nelleke Aiello, for debugging of Tutorial mode, Donald A. Norman for several useful suggestions on an earlier version of the paper, and especially the late Jaime R. Carbonell for his many ideas, in addition to the SCHOLAR program.

2. There are two kinds of topics: attributes, such as Names-and-Locations-of-Countries or Geographical-Features, and values, such as South America or Cape Horn. We have adopted the convention of hyphenating an attribute when it appears as a topic to indicate that it is to be read as a unit. A subtopic is never simply Geographical-Features or Cape Horn but the entire embedded chain up to the main topic (e.g. Cape Horn as a Geographical-Feature of South America).

3. The comments in parentheses are by the authors. The dialogue is verbatim except for a few minor grammatical changes to make the text into intelligible sentences.

4. In SCHOLAR's data base there are embedded attributes, as for example Bordering-Countries is embedded under Location in Figure 1. For such cases the rules must be modified as follows: Rule 1 selects an entire embedded chain of attributes plus the associated value. Rule 4 adds the value and each of the embedded attributes separately to the pushdown list of topics. Thus Northern Bordering-Countries might be followed by Western Bordering-Countries as a topic, when the embedded attribute Northern is replaced by Western. Such embedding of attributes also occurred with the tutors. For example, for the first tutor in the dialogue shown, the subattributes Rivers and Mountains were apparently embedded under Geographical-Features of South America, and the subattributes Religion and Language under Population of Argentina (see Figure 4).

5. The letters in parenthesis denote how each piece of information was brought up (see text). Q denotes a question by the tutor that the student answered correctly. P denotes a presentation by the tutor without any preceding question. QP denotes a question by the tutor, which the student failed to answer, and so the tutor presented the correct answer. S denotes the occasional cases where the student volunteered a piece of information.

6. If the tutor had been teaching the basic concepts of geography rather than facts about South American geography, then presumably the questioning about these basic concepts would have been tied to the structure of the discussion.

7. This module was designed and implemented by Nelleke Aiello, with the help of Susan Graesser.

8. We would like to thank Marshall Farr, our contract monitor at ONR, who suggested this idea to us.

9. The approach described here comes down to us from Ross Quillian, and through him from Allen Newell and Herbert Simon.

DISTRIBUTION LIST

4 Dr. Marshall J. Farr, Director
Personnel & Training
Research Programs
Office of Naval Research
Arlington, VA 22217

1 Director
ONR Branch Office
495 Summer Street
Boston, MA 02210
ATTN: C. M. Harsh

1 Director
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101
ATTN: E. E. Gloye

1 Director
ONR Branch Office
536 South Clark Street
Chicago, IL 60605
ATTN: M. A. Bertin

1 Office of Naval Research
Area Office
207 West 24th Street
New York, NY 10011

2 Director
Advanced Research Projects
Agency
1400 Wilson Boulevard
Arlington, VA 22209
Attention: Program Management
DODAAD Code
HX1241

1 STOIAC
Battelle Memorial Institute
505 King Avenue
Columbus, Ohio 43201
Attention: H8-1 Code
79986

6 Director
Naval Research Laboratory
Code 2627
Washington, DC 20390

12 Defense Documentation
Center
Cameron Station,
Building 5
5010 Duke Street
Alexandria, VA 22314

1 Chairman
Behavioral Science Department
Naval Command and Management
Division
U.S. Naval Academy
Luce Hall
Annapolis, MD 21402

1 Chief of Naval Technical
Training
Naval Air Station Memphis(75)
Millington, TN 38054
ATTN: Dr. G. D. Mayo

1 Chief of Naval Training
Naval Air Station
Pensacola, FL 32508
ATTN: CAPT Bruce Stone, USN

1 TACTEC
Battelle Memorial Institute
505 King Avenue
Columbus, Ohio 43201
Attention H8-1 Code
79986

1 Defense Contract Administration
Services Region
666 Summer Street
Boston, MA 02210
Attention: DODAAD Code
52202A

6 Director
Naval Research Laboratory
Washington, D.C. 20375
Attention: DODAAD Code
N00173
Code 2629

1 Program Coordinator
Bureau of Medicine and Surgery
Code 71G
Department of the Navy
Washington, D.C. 20372

1 Commanding Officer
Naval Medical Neuropsychiatric
Research Unit
San Diego, CA 92152

1 Dr. Harold Booher
NAVAIR 415C
Naval Air Systems Command
5600 Columbia Pike
Falls Church, VA 22042

1 CAPT John F. Riley, USN
Command Officer
U.S. Naval Amphibious School
Coronado, CA 92155

1 Special Assistant for Manpower
OASN (M&RA)
The Pentagon, Room 4E794
Washington, D.C. 20350

1 CDR Richard L. Martin, USN
COMFAIRMIRAMAR F-14
NAS Miramar, CA 92145

1 Research Director, Code 06
Research and Evaluation
Department
U.S. Naval Examining Center
Great Lakes, IL 60088
ATTN: C. S. Winiewicz

1 Commander
Naval Air Reserve
Naval Air Station
Glenview, IL 60026

1 Commander
Naval Air Systems Command
Department of the Navy
AIR-413C
Washington, D.C. 20360

1 Mr. Lee Miller (AIR 413E)
Naval Air Systems Command
5600 Columbia Pike
Falls Church, VA 22042

1 Dr. John J. Collins
Chief of Naval Operations
(OP-987F)
Department of the Navy
Washington, D.C. 20350

1 Technical Library (Pers-11B)
Bureau of Naval Personnel
Department of the Navy
Washington, D.C. 20360

1 Dr. Richard J. Niehaus
Office of Civilian
Manpower Management
Code 06A
Department of the Navy
Washington, D.C. 20390

1 Dr. James J. Regan
Technical Director
Navy Personnel Research
and Development Center
San Diego, CA 92152

1 Chief
Bureau of Medicine and Surgery
Code 413
Washington, D.C. 20372

1 Mr. Arnold Rubinstein
Naval Material Command
(NMAT-03424)
Room 820, Crystal Plaza #6
Washington, D.C. 20360

1 Superintendent
Naval Postgraduate School
Monterey, CA 92940
ATTN: Library (Code 2124)

1 Technical Library
Naval Ship Systems Command
National Center, Building 3
Room 3S08
Washington, D.C. 20360

1 Chief of Naval Training Support
Code N-21, Building 45
Naval Air Station
Pensacola, FL 32508

1 Dr. H. Wallace Sinaiko
c/o Office of Naval Research
(Code 450)
Psychological Sciences Division
Arlington, VA 22217

1 Commandant
U. S. Army Institute of
Administration
ATTN: EA
Fort Benjamin Harrison, IN 46216

1 Armed Forces Staff College
Norfolk, VA 23511
ATTN: Library

1 Commander
U. S. Theatre Army Support
Command, Europe
ATTN: Asst. DCSPER (Education)
APO New York 09058

1 Commanding Officer
Naval Personnel Research
and Development Center
San Diego, CA 92152

1 Mr. George N. Graine
Naval Ship Systems Command
(SHIPS 03H)
Department of the Navy
Washington, D.C. 20360

1 Commanding Officer
Service School Command
U.S. Naval Training Center
San Diego, CA 92133
ATTN: Code 303

1 Dr. William L. Maloy
Principal Civilian Advisor
for Education and Training
Naval Training Command,
Code 01A
Pensacola, FL 32508

1 Dr. Martin F. Wiskoff
Navy Personnel Research
and Development Center
San Diego, CA 92152

1 Chief, Unit Training and
Educational Technology
Systems
U. S. Army Research
Institute for the Behavioral
and Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

1 Director of Research
U. S. Army Armor Human
Research Unit
ATTN: Library
Building 2422 Morade Street
Fort Knox, KY 40121

1 U. S. Army Research Institute
for the Behavioral and
Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

1 Commanding Officer
ATTN: LTC Montgomery
USADC - PASA
Ft. Benjamin Harrison, IN 46249

1 Dr. John L. Kobrick
Military Stress Laboratory
U. S. Army Research Institute
of Environmental Medicine
Natick, MA 01760

1 Commandant
United States Army
Infantry School
ATTN: ATSIN-H
Fort Benning, GA 31905

1 U. S. Army Research Institute
Commonwealth Building
Room 239
1300 Wilson Boulevard
Arlington, VA 22209
ATTN: Dr. R. Dusek

1 Mr. Edmund F. Fuchs
U. S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209

1 Dr. Stanley L. Cohen
Work Unit Area Leader
Organizational Development
Work Unit
Army Research Institute
for the Behavioral and
Social Sciences
1300 Wilson Boulevard
Arlington, VA 22209

1 Dr. Leon H. Nawrocki
U. S. Army Research Institute
Rosslyn Commonwealth Building
1300 Wilson Boulevard
Arlington, VA 22209

1 Dr. Martin Rockway
Technical Training Division
Lowry Air Force Base
Denver, CO 80230

1 Maj. P. J. DeLeo
Instructional Technology Branch
AF Human Resources Laboratory
Lowry AFB, CO 80230

1 Headquarters, U.S. Air Force
Chief, Personnel Research
and Analysis Division
(AF/DPSY)
Washington, DC 20330

1 Research and Analysis Division
AF/DPXYR Room 4C200
Washington, DC 20330

1 AFHRL/AS (Dr. G. A. Eckstrand
Wright-Patterson Air Force Base
Ohio 45433

1 AFHRL (AST/Dr. Ross L. Morgan
Wright Patterson Air Force Base)
Ohio 45433

1 AFHRL/MD
701 Prince Street
Room 200
Alexandria, VA 22314

1 AFOSR(NL)
1400 Wilson Boulevard
Arlington, VA 22209

1 Commandant
USAF School of Aerospace
Medicine
Aeromedical Library (SUL-4)
Brooks AFB, TX 78235

1 CAPT Jack Thorpe, USAF
Department of Psychology
Bowling Green State University
Bowling Green, OH 43403

1 Headquarters Electronic
Systems Division
ATTN: Dr. Sylvia R. Mayer/MCIT
LG Hanscom Field
Bedford, MA 01730

1 Commandant, Marine Corps
Code AO1M-2
Washington, D.C. 20380

1 COL George Caridakis
Director, Office of
Manpower Utilization
Headquarters, Mareine Corps
(A01H), MCB
Quantico, VA 22134

1 Dr. A. K. Slafkosky
Scientific Advisor (Code Ax)
Commandant of the Marine Corps
Washington, DC 20380

1 Mr. E. A. Dover
Manpower Measurement Unit
(Code AOIM-2)
Arlington Annex, Room 2413
Arlington, VA 20370

1 Mr. Joseph J. Cowan, Chief
Psychological Research
Branch (P-1)
U. S. Coast Guard Headquarters
400 Seventh Street, SW
Washington, DC 20590

1 Dr. John Ford, Jr.
Navy Personnel Research
and Development Center
San Diego, CA 92152

1 LCDR Charles Theisen, Jr.,
MSC, USN, 4024
Naval Air Development Center
Warminster, PA 18974

1 Mr. Helga Yeich, Director
Program Management, Defense
Advanced Research
Projects Agency
1400 Wilson Boulevard
Arlington, VA 22209

1 Mr. William J. Stormer
DOD Computer Institute
Washington Navy Yard
Building 175
Washington, DC 20374

1 Dr. Ralph R. Canter
Director for Manpower Research
Office of Secretary of Defense
The Pentagon, Room 3C980
Washington, DC 20301

1 Office of Computer Information
Institute for Computer Sciences
and Technology
National Bureau of Standards
Washington, DC 20234

1 Dr. Eric McWilliams
Program Manager
Technology and Systems, TIE
National Science Foundation
Washington, DC 20550

1 Dr. Scarvia Anderson
Executive Director for Special Development
Educational Testing Service
Princeton, NJ 08540

1 Dr. Richard C. Atkinson
Stanford University
Department of Psychology
Stanford, CA 94305

1 Dr. Bernard M. Bass
University of Rochester
Management Research Center
Rochester, NY 14627

1 Mr. Edmund C. Berkeley
Berkeley Enterprises, Inc.
815 Washington Street
Newtonville, MA 02160

1 Dr. David G. Bowers
University of Michigan
Institute for Social Research
P. O. Box 1248
Ann Arbor, Michigan 48106

1 Dr. Ronald P. Carver
American Institutes for Research
8555 Sixteenth Street
Silver Spring, MD 20910

1 Century Research Corporation
4113 Lee Highway
Arlington, VA 22207

1 Dr. Robert Glaser, Director
University of Pittsburgh
Learning Research and Development Center
Pittsburgh, PA 15213

1 Dr. Kenneth E. Clark
University of Rochester
College of Arts and Sciences
River Campus Station
Rochester, NY 14627

1 Dr. Rene' V. Dawis
University of Minnesota
Department of Psychology
Minneapolis, MN 55455

2 ERIC
Processing and Reference Facility
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Victor Fields
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Mr. H. Dean Brown
Stanford Research Institute
333 Ravenswood Avenue
Menlo Park, CA 94025

1 Mr. Michael W. Brown
Operations Research, Inc.
1400 Spring Street
Silver Spring, MD 20910

1 Dr. Edwin A. Fleishman
American Institutes for Research
8555 Sixteenth Street
Silver Spring, MD 20910

1 Dr. Albert S. Glickman
American Institutes for Research
8555 Sixteenth Street
Silver Spring, MD 20910

1 Dr. Duncan N. Hansen
Florida State University
Center for Computer-Assisted
Instruction
Tallahassee, FL 32306

1 Dr. Henry J. Hamburger
University of California
School of Social Sciences
Irvine, CA 92664

1 Dr. Richard S. Hatch
Decision Systems Associates Inc.
11428 Rockville Pike
Rockville, MD 20852

1 Dr. M. D. Havron
Human Sciences Research, Inc.
Westgate Industrial Park
7710 Old Springhouse Road
McLean, VA 22101

1 Human Resources Research
Organization
Division #3
P. O. Box 5787
Presidio of Monterey, CA 93940

1 Human Resources Research
Organization
Division #4, Infantry
P. O. Box 2086
Fort Benning, GA 31905

1 Human Resources Research
Organization
Division #5, Air Defense
P. O. Box 6057
Fort Bliss, TX 79916

1 Human Resources Research
Organization
Division #6, Library
P. O. Box 428
Fort Rucker, AL 36360

1 Dr. Lawrence B. Johnson
Lawrence Johnson and
Associates, Inc.
200 S Street, NW, Suite 502
Washington, DC 20009

1 Dr. Norman J. Johnson
Carnegie-Mellon University
Graduate School of
Industrial Admin.
Pittsburgh, PA 15213

1 Dr. David Klahr
Carnegie-Mellon University
Graduate School of
Industrial Admin.
Pittsburgh, PA 15213

1 Dr. Robert R. Mackie
Human Factors Research, Inc.
6780 Cortona Drive
Santa Barbara Research Park
Goleta, CA 93017

1 Dr. Andrew R. Molnar
Technological Innovations
in Education
National Science Foundation
Washington, DC 20550

1 Dr. Leo Munday
Vice President
American College Testing
Program
P.O. Box 168
Iowa City, IA 52240

1 Dr. Donald A. Norman
University of California,
San Diego
Center for Human Information
Processing
La Jolla, CA 92037

1 Mr. Luigi Petrullo
2431 North Edgewood Street
Arlington, VA 22207

1 Dr. Robert D. Pritchard
Assistant Professor of Psychology
Purdue University
Lafaytte, IN 47907

1 Dr. Diane M. Ramsey-Klee
R-K Research & System Design
3947 Ridgemont Drive
Malibu, CA 90265

1 Dr. Joseph W. Rigney
Behavioral Technology
Laboratories
University of Southern
California
3717 South Grand
Los Angeles, CA 90007

1 Dr. Leonard L. Rosenbaum, Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. George E. Rowland
Rowland and Company, Inc.
P. O. Box 61
Haddonfield, NJ 08033

1 Dr. Arthur I. Siegel
Applied Psychological
Services and Science Center
404 East Lancaster Avenue
Wayne, PA 19087

1 Mr. Dennis J. Sullivan
725 Benson Way
Thousand Oaks, CA 91360

1 Mr. Dexter Fletcher
Department of Psychology
University of Illinois
at Chicago Circle
Box 4348
Chicago, IL 60680

1 Dr. Benton J. Underwood
Northwestern University
Department of Psychology
Evanston, IL 60201

1 Dr. David J. Weiss
University of Minnesota
Department of Psychology
Minneapolis, MN 55455

1
Denver Research Institute
University of Denver
Denver, CO 80210

1 Dr. Kenneth Wexler
University of California
School of Social Sciences
Irvine, CA 92664

1 Dr. John Annett
The Open University
Milton Keynes
Buckinghamshire
ENGLAND

1 Dr. Milton S. Katz
MITRE Corporation
Westgate Research Center
McLean, VA 22101

1 Dr. Charles A. Ullmann
Director, Behavioral
Sciences Studies
Information Concepts
Incorporated
1701 No. Ft. Myer Drive
Arlington, VA 22209

1 Dr. Anita West
Denver Research Institute
University of Denver
Denver, CO 802]0